AMERICAN JOURNAL OF MATHEMATICS

FOUNDED BY THE JOHNS HOPKINS UNIVERSITY

PUBLISHED UNDER THE JOINT AUSPICES OF

THE JOHNS HOPKINS UNIVERSITY
AND
THE AMERICAN MATHEMATICAL SOCIETY

VOLUME LXXVIII

1956

THE JOHNS HOPKINS PRESS
BALTIMORE 18, MARYLAND
U. S. A.

PRINTED IN THE UNITED STATES OF AMERICA
BY J. H. FURST COMPANY, BALTIMORE, MARYLAND

INDEX

REPRESENTATIONS OF SEMISIMPLE LIE GROUPS, V.*

By HARISH-CHANDRA.

1. Introduction. Let G be a connected semisimple Lie group and A a Cartan subgroup of G. Under the assumption that the image of A in the adjoint group of G is compact, we have studied in detail [5(f)] certain irreducible representations of the Lie algebra of G and seen that they can all be 'extended" to representations of the group (Theorem 4 of [5(f)]). In the present paper we shall obtain them directly as irreducible representations of G on certain Hilbert spaces consisting of holomorphic functions on a suitable complex manifold.

It turns out that these representations are very intimately related with the finite-dimensional ones (Lemma 14) and the two have some striking similarities (Lemmas 6 and 12). Moreover, as we shall see in Theorem 3, under appropriate conditions some of these representations are unitary. In the last section we obtain a result on their characters. Some of the deeper analogies between these representations of G and those of a compact smisimple group (see [5(e)]) will be discussed in another paper.

2. Certain complex manifolds. We keep to the notation of [5(f), § 3]. Let $\mathfrak{g}_0$ be a semisimple Lie algebra over the field R of real numbers. Define $\mathfrak{k}_0$ and $\mathfrak{p}_0$ as in [5(b), § 2] and let $\mathfrak{h}_0$ be a maximal abelian subalgebra of $\mathfrak{k}_0$. *We shall assume that $\mathfrak{h}_0$ is also maximal abelian in $\mathfrak{g}_0$* (see [5(f), § 3]. Let C denote the field of complex numbers. We complexify $\mathfrak{g}_0, \mathfrak{h}_0, \mathfrak{k}_0, \mathfrak{p}_0$ to $\mathfrak{g}, \mathfrak{h}, \mathfrak{k}, \mathfrak{p}$ respectively. Let $X \rightarrow adX$ $(X \varepsilon \mathfrak{g})$ denote the adjoint representation of $\mathfrak{g}$ and θ the automorphism of order 2 given by $\theta(X+Y)=X-Y$ $(X \varepsilon \mathfrak{k}, Y \varepsilon \mathfrak{p})$. Put[1] $\mathfrak{u}=\mathfrak{k}_0+(-1)^{\frac{1}{2}}\mathfrak{p}_0$. Then $\mathfrak{u}$ is a compact real form of $\mathfrak{g}$. We denote by $\tilde{\theta}$ and η the conjugations of $\mathfrak{g}$ with respect to $\mathfrak{u}$ and $\mathfrak{g}_0$ respectively. Then

$$\tilde{\theta}(X+(-1)^{\frac{1}{2}}Y)=X-(-1)^{\frac{1}{2}}Y \quad (X, Y \varepsilon \mathfrak{u}),$$
$$\eta(X'+(-1)^{\frac{1}{2}}Y')=X'-(-1)^{\frac{1}{2}}Y' \quad (X', Y' \varepsilon \mathfrak{g}_0)$$

and $\eta=\tilde{\theta}\theta$.

We consider an arbitrary but fixed order (see [5(f), § 2]) in the space

* Received July 8, 1955.

[1] We fix once for all a square-root of -1 in C and denote it by $(-1)^{\frac{1}{2}}$.

$\mathfrak{F}_R$ of real linear functions [2] on $\mathfrak{h}$. Put [3] $\mathfrak{n}_+ = \sum_{\alpha>0} CX_\alpha$ and $\mathfrak{n}_- = \sum_{\alpha>0} CX_{-\alpha}$ (where α runs over all positive roots) and let G_c be the simply connected complex analytic group with the Lie algebra $\mathfrak{g}$. We denote by G_0, K_0, A_0, A_-, A_c, N_c^+, N_c^-, U the (real) analytic subgroups of G_c corresponding to $\mathfrak{g}_0$, $\mathfrak{k}_0$, $\mathfrak{h}_0$, $(-1)^{\frac{1}{2}}\mathfrak{h}_0$, $\mathfrak{h}$, $\mathfrak{n}_+$, $\mathfrak{n}_-$, $\mathfrak{u}$ respectively. Then K_0, A_0 and U are compact and U, A_+, N_c^+, N_c^- are simply connected.[4] We now 'extend' θ, $\tilde{\theta}$ and η to automorphisms of G_c. Since $\tilde{\theta}(\mathfrak{n}_+) = \eta(\mathfrak{n}_+) = \mathfrak{n}_-$, it follows that $\tilde{\theta}(N_c^+)$ $\eta(N_c^+) = N_c^-$. According to a theorem of Iwasawa [6] (see also Lemma 26 of [5(b)]) the mapping $(u, h, n) \to uhn$ $(u \varepsilon U, h \varepsilon A_+, n \varepsilon N_c^+)$ is a one-one regular analytic mapping of $U \times A_+ \times N_c^+$ onto G_c. We denote by $z \to Ad(z)$ $(z \varepsilon G_c)$ the adjoint representation of G_c.

LEMMA 1. *$N_c^- \cap A_+N_c^+ = \{1\}$ and $N_c^-A_+ \cap N_c^+ = \{1\}$.*

For suppose $an \varepsilon N_c^-$ $(a \varepsilon A_+, n \varepsilon N_c^+)$. Since $\mathfrak{n}_+$ is a nilpotent algebra, it is mapped onto N_c^+ under the exponential mapping (see Birkhoff [1]). Similarly for $\mathfrak{n}_-$. Hence we can choose $X \varepsilon \mathfrak{n}_+$ and $Y \varepsilon \mathfrak{n}_-$ such that $n = \exp X$, $an = \exp Y$. Now suppose $an \neq 1$. Then $Y \neq 0$ and there exists an element $H \varepsilon \mathfrak{h}$ such that $\exp(adY)H = H + Z$ where $Z \varepsilon \mathfrak{n}_-$ and $Z \neq 0$ (Lemma 8 of [5(d)]). Therefore it is obvious that

$$H + Z = \exp(adY)H = Ad(an)H = H + Z',$$

where $Z' \varepsilon \mathfrak{n}_+$. But since $\mathfrak{n}_+ \cap \mathfrak{n}_- = \{0\}$ it follows that $Z = Z' = 0$ and so we get a contradiction. This proves that $N_c^- \cap A_+N_c^+ = \{1\}$. If we transform this result under the mapping $z \to \tilde{\theta}(z^{-1})$ $(z \varepsilon G_c)$, we find that $N_c^-A_+ \cap N_c^+ = \{1\}$.

COROLLARY. *$G_0 \cap A_+N_c^+ = \{1\}$ and $N_c^-A_+ \cap G_0 = \{1\}$.*

For suppose $x = an$ $(x \varepsilon G_0, a \varepsilon A_+, n \varepsilon N_c^+)$. Then $an = x = \eta(x) = a^{-1}\eta(n)$, since $\eta(H) = -H$ if $H \varepsilon (-1)^{\frac{1}{2}}\mathfrak{h}_0$. Therefore

$$\eta(n) = a^2 n \varepsilon N_c^- \cap A_+N_c^+ = \{1\}$$

and so $a^2 = n = 1$. Since A_+ is abelian and simply connected this implies that $a = 1$ and therefore $x = 1$. The second assertion follows from the first under the mapping $z \to \eta(z^{-1})$ $(z \varepsilon G_c)$.

[2] A linear function on $\mathfrak{h}$ is called real if it takes real values on $(-1)^{\frac{1}{2}}\mathfrak{h}_0$ (see [5(f), § 2]).

[3] Any undefined terms or symbols should automatically be given the same meaning as in [5(f)].

[4] The proofs of these statements are well known and the necessary references can be found in [5(b)]. See in particular [3] and [7].

LEMMA 2. *$G_0A_+N_c^+$ and $N_c^-A_+G_0$ are open in G_c.*

Since $N_c^-A_+G_0$ is the image of $G_0A_+N_c^+$ under the topological mapping $z \to \eta(z^{-1})$ $(z \in G_c)$ it is enough to prove that $G_0A_+N_c^+$ is open. Let $\mathfrak{s}_+ = (-1)^{\frac{1}{2}}\mathfrak{h}_0 + \mathfrak{n}_+$. Then $\mathfrak{s}_+$ is the Lie algebra of $A_+N_c^+$ and so in view of the corollary $\mathfrak{g}_0 \cap \mathfrak{s}_+ = \{0\}$. Hence

$$\dim_R(\mathfrak{g}_0 + \mathfrak{s}_+) = \dim_R \mathfrak{g}_0 + \dim_R \mathfrak{s}_+.$$

On the other hand $\mathfrak{g}$ is the direct sum of $\mathfrak{h}$, $\mathfrak{n}_+$ and $\mathfrak{n}_-$ and therefore

$$\begin{aligned}\dim_R \mathfrak{g}_0 &= \dim_C \mathfrak{g} = \dim_C \mathfrak{h} + \dim_C \mathfrak{n}_+ + \dim_C \mathfrak{n}_- \\ &= \dim_R \mathfrak{h}_0 + 2\dim_C \mathfrak{n}_+ = \dim_R \mathfrak{h}_0 + \dim_R \mathfrak{n}_+ = \dim_R \mathfrak{s}_+.\end{aligned}$$

This shows that $\dim_R(\mathfrak{g}_0 + \mathfrak{s}_+) = 2\dim_R \mathfrak{g}_0 = \dim_R \mathfrak{g}$, and therefore $\mathfrak{g}$ is the direct sum of $\mathfrak{g}_0$ and $\mathfrak{s}_+$. Our assertion now follows from Lemma 26 of [5(b)].

Since A_+ is simply connected, for every $a \in A$ there exists a unique element $H \in (-1)^{\frac{1}{2}}\mathfrak{h}_0$ such that $a = \exp H$. We denote this element by $\log a$. Also any element $z \in G_c$ can be written uniquely in the form $z = uhn$ $(u \in U$, $h \in A_+$, $n \in N_c^+)$. We write $u(z)$ and $H(z)$ to denote u and $\log h$ respectively. Then $z \to u(z)$ and $z \to H(z)$ are (real) analytic mappings of G_c into U and $(-1)^{\frac{1}{2}}\mathfrak{h}_0$ respectively.

LEMMA 3. *Let 2ρ denote the sum of all the positive roots of $\mathfrak{g}$. Then if dx is the Haar measure of G_0,*

$$\int_{G_0} e^{-4\rho(H(x))}dx < \infty.$$

Let ψ denote the mapping $x \to u(x)$ $(x \in G_0)$ of G_0 into U. It follows from the corollary to Lemma 1 that ψ is univalent. Let $x = us$ $(x \in G_0, u \in U$, $s \in A_+N_c^+)$. Then a simple calculation shows that if $X \in \mathfrak{g}_0$, $(d\psi)_x X = V$ where V is the element of $\mathfrak{u}$ determined by the condition $V \equiv Ad(s)X \bmod \mathfrak{s}_+$. (Here $(d\psi)_x$ is the differential of ψ at x (see Chevalley [4]). Now we regard $\mathfrak{g}/\mathfrak{s}_+$ as a vector space over R and denote by D its endomorphism corresponding to $Ad(s)$. Since $\det(Ad(s)) = 1$, $\det D = \det(Ad(s^{-1}))_{\mathfrak{s}_+}$ where $(Ad(s^{-1}))_{\mathfrak{s}_+}$ is the restriction of $Ad(s^{-1})$ on $\mathfrak{s}_+$. Then if $s = an$ $(a \in A_+, n \in N_c^+)$,

$$\det(Ad(s^{-1}))_{\mathfrak{s}_+} = |e^{-2\rho(\log a)}|^2 = e^{-4\rho(\log a)}$$

and therefore $\det D = e^{-4\rho(\log a)} = e^{-4\rho(H(x))}$. This calculation shows that

$$du = e^{-4\rho(H(x))}dx,$$

where du and dx are the Haar measures on U and G_0 respectively. Hence

$$\int_{G_0} e^{-4\rho(H(x))}dx = \int_{\psi(G_0)} du \leqq \int_U du < \infty$$

because U is compact.

Let Q_+ be the set of all totally positive roots of $\mathfrak{g}$ and Q the remaining set of positive roots (see [5(f)]). Put $\mathfrak{p}_+ = \sum_{\beta \varepsilon Q_+} CX_\beta$, $\mathfrak{p}_- = \sum_{\beta \varepsilon Q_+} CX_{-\beta}$, $\mathfrak{p}' = \sum_{\beta \varepsilon Q} (CX_\beta + CX_{-\beta})$ and $\mathfrak{m} = \mathfrak{k} + \mathfrak{p}'$. Then $\mathfrak{p}_+$, $\mathfrak{p}_-$, $\mathfrak{m}$ are all subalgebras of $\mathfrak{g}$ which is a direct sum of these three. Moreover $\mathfrak{p}_+$ and $\mathfrak{p}_-$ are abelian and $[\mathfrak{m}, \mathfrak{p}_+] \subset \mathfrak{p}_+$, $[\mathfrak{m}, \mathfrak{p}_-] \subset \mathfrak{p}_-$ (see [5(f), §§4 and 5] for the proofs of these statements). Let P_c^+, P_c^- and M_c be the complex analytic subgroups of G_c corresponding to $\mathfrak{p}_+$, $\mathfrak{p}_-$ and $\mathfrak{m}$ respectively.

LEMMA 4. *The mapping $(q, m, p) \to qmp$ $(q \varepsilon P_c^-, m \varepsilon M_c, p \varepsilon P_c^+)$ of $P_c^- \times M_c \times P_c^+$ into G_c is univalent, holomorphic and regular.*

First we claim that $P_c^- M_c \cap P_c^+ = \{1\}$. For suppose $y \varepsilon P_c^- M_c \cap P_c^+$. Since P_c^+ is abelian, we can choose $Y \varepsilon \mathfrak{p}_+$ such that $y = \exp Y$. Also since $[\mathfrak{m}, \mathfrak{p}_-] \subset \mathfrak{p}_-$ and since $y \varepsilon P_c^- M_c$, it is clear that $Ad(y)\mathfrak{p}_- = \mathfrak{p}_-$. Now suppose $y \neq 1$ so that $Y \neq 0$. Then $Y = \sum_\gamma c_\gamma X_\gamma$ where γ runs over all the totally positive roots and $c_\gamma \varepsilon C$. Let γ_0 be the lowest root such that $c_{\gamma_0} \neq 0$. Then $[Y, X_{-\gamma_0}] \equiv c_{\gamma_0} H_{\gamma_0} \bmod \mathfrak{n}_+$ and therefore it is obvious that

$$Ad(y) X_{-\gamma_0} \equiv X_{-\gamma_0} + c_{\gamma_0} H_{\gamma_0} \bmod \mathfrak{n}_+.$$

However $H_{\gamma_0} \notin \mathfrak{n}_+ + \mathfrak{n}_-$ and so it follows that $Ad(y) X_{-\gamma_0} \notin \mathfrak{p}_-$. Since this contradicts the fact that $Ad(y)\mathfrak{p}_- = \mathfrak{p}_-$ we must have $y = 1$ and therefore $P_c^- M_c \cap P_c^+ = \{1\}$. Taking the image of this equality under the mapping $z \to \eta(z^{-1})$ $(z \varepsilon G_c)$ we get $P_c^- \cap M_c P_c^+ = \{1\}$. Now if we make use of the fact that both $P_c^- M_c$ and $M_c P_c^+$ are subgroups of G_c, the univalence of our mapping follows straightaway. That it is holomorphic is an immediate consequence of the complex analyticity of G_c. Similarly the regularity follows from the relation $\mathfrak{g} = \mathfrak{p}_- + \mathfrak{m} + \mathfrak{p}_+$ (Lemma 26 of [5(b)]).

It is clear from Lemmas 2 and 4 that $G_0 A_+ N_c^+$, $N_c^- A_+ G_0$ and $P_c^- M_c P_c^+$ are open connected subsets of G_c and therefore they may be regarded as open submanifolds of G_c.

LEMMA 5. *$N_c^- A_+ G_0 A_+ N_c^+$ is contained in $P_c^- M_c P_c^+$.*

Let $\exp \mathfrak{p}_0$ denote the set of all elements in G_0 of the form $\exp X$ $(X \varepsilon \mathfrak{p}_0)$. Then if $p \varepsilon \exp \mathfrak{p}_0$ and $p = uan$ $(u \varepsilon U,\ a \varepsilon A_+,\ n \varepsilon N_c^+)$,

$$p^{-1} = \theta(p) = ua^{-1}\tilde{\theta}(n)$$

since $\bar{\theta}(u)=u$ and $\bar{\theta}(a)=a^{-1}$. Therefore

$$p^2=\theta(n^{-1})a^2n\,\varepsilon\,N_c^-A_+N_c^+.$$

Now let X be any element in $\mathfrak{p}_0$ and put $p=\exp(\frac{1}{2}X)$. Then

$$p^2=\exp X\,\varepsilon\,N_c^-A_+N_c^+$$

and therefore $\exp\mathfrak{p}_0\subset N_c^-A_+N_c^+$. On the other hand $\mathfrak{m}+\mathfrak{p}_+\supset\mathfrak{h}+\mathfrak{n}_+$ and $\mathfrak{p}_-+\mathfrak{m}\supset\mathfrak{n}_-+\mathfrak{h}$. Therefore $A_+N_c^+\subset M_cP_c^+$ and $N_c^-A_+\subset P_c^-M_c$ and so it follows that $\exp\mathfrak{p}_0\subset P_c^-M_cP_c^+$. But $G_0=K_0(\exp\mathfrak{p}_0)$ (see Cartan [3] and Mostow [7]) and $K_0P_c^-=P_c^-K_0$ since $[\mathfrak{k},\mathfrak{p}_-]\subset\mathfrak{p}_-$. Moreover $K_0\subset M_c$ and therefore $G_0\subset P_c^-M_cP_c^+$. However we have seen already that $A_+N_c^+\subset M_cP_c^+$ and $N_c^-A_+\subset P_c^-M_c$ and so it follows that $N_c^-A_+G_0A_+N_c^+\subset P_c^-M_cP_c^+$.

3. The simply connected covering manifold of $N_c^-A_+G_0$. Let S denote the subgroup $N_c^-A_+$ of G_c. Then we have seen that $\bar{W}=SG_0$ is an open submanifold of G_c. Since $\mathfrak{g}=\mathfrak{n}_-+(-1)^{\frac{1}{2}}\mathfrak{h}_0+\mathfrak{g}_0$ (see the proof of Lemma 2) the maping $(s,\bar{x})\to s\bar{x}$ $(s\,\varepsilon\,S,\bar{x}\,\varepsilon\,G_0)$ of $S\times G_0$ into $\bar{W}$ is everywhere regular (Lemma 26 of [5(b)]) and therefore open. Also $S\cap G_0=\{1\}$ (corollary to Lemma 1) and so this mapping is univalent. Hence it is a homeomorphism. Let G be the simply connected covering group of G_0 and let $x\to\bar{x}$ $(x\,\varepsilon\,G)$ denote the natural homomorphism of G onto G_0. Since S is simply connected, $S\times G$ is a simply connected covering space of $S\times G_0$. Therefore we may also regard it as a covering space of $\bar{W}$ under the mapping $\nu:(s,x)\to s\bar{x}$ $(s\,\varepsilon\,S,x\,\varepsilon\,G)$. Since $\bar{W}$ is a complex manifold, we can introduce a complex structure in $S\times G$ in such a way that it becomes a covering manifold of $\bar{W}$ with respect to the mapping ν. Let W denote the complex manifold arising from $S\times G$ in this way. We identify S and G with subsets of W under the topological mappings $s\to(s,1)$ $(s\,\varepsilon\,S)$ and $x\to(1,x)$ $(x\,\varepsilon\,G)$. Then $S\cap G=(1,1)$ is the common unit element of S and G which we shall denote by 1. Let $\bar{W}_l$ and $\bar{W}_r$ respectively be the set of all elements $\bar{v}$ and $\bar{w}$ in $\bar{W}$ such that $\bar{v}\bar{W}\subset\bar{W}$ and $\bar{W}\bar{w}\subset\bar{W}$. Put $W_l=\nu^{-1}(\bar{W}_l)$ and $W_r=\nu^{-1}(\bar{W}_r)$. Then if $z\,\varepsilon\,W_l$, the mapping $\bar{l}_z:\bar{w}\to\nu(z)\bar{w}$ $(\bar{w}\,\varepsilon\,\bar{W})$ is obviously holomorphic on $\bar{W}$. Hence it is clear that there exists exactly one holomorphic mapping l_z of W into itself such that $\nu\circ l_z=\bar{l}_z\circ\nu$ and $l_z(1)=z$. Similarly if $z\,\varepsilon\,W_r$, there exists just one holomorphic mapping r_z of W into itself such that $\nu(r_z(w))=\nu(w)\nu(z)$ $(w\,\varepsilon\,W)$ and $r_z(1)=z$. For convenience we shall write l_zw and r_vw instead of $l_z(w)$ and $r_v(w)$ respectively $(z\,\varepsilon\,W_l,v\,\varepsilon\,W_r,w\,\varepsilon\,W)$. Let A be the analytic subgroup of G corresponding to $\mathfrak{h}_0$. Then it is obvious that $S\subset W_l$, $G\subset W_r$ and $A\subset W_l\cap W_r$. Moreover if $u,v\,\varepsilon\,W_l$, $z=l_uv$ also

lies in W_l and $l_z = l_u l_v$. This shows that the multiplication in W_l, defined by the rule $uv = l_u v$ $(u, v \in W_l)$, is associative. Similarly we define an associative multiplication in W_r by $zw = r_w z$ $(z, w \in W_r)$. It is easy to check that these two multiplications coincide on $W_r \cap W_l$.

We recall that A_c is the complex analytic subgroup of G_c corresponding to $\mathfrak{h}$. Put $\tilde{A}_c = \nu^{-1}(A_c)$. Since A contains the center[5] of G, $\tilde{A}_c = A_+ \times A$ and so it is connected. Also $A_c \subset \bar{W}_l$ and therefore $\tilde{A}_c \subset W_l$. It is easy to verify that $\tilde{A}_c$ is a group (with respect to the multiplication defined in W_l) and actually it can be regarded as a covering group of A_c under the mapping $a \to \nu(a)$ $(a \in \tilde{A}_c)$. Since $N_c^- \subset W_l$, the product na $(n \in N_c^-, a \in \tilde{A}_c)$ is well defined in W_l.

4. Holomorphic functions on W. By a holomorphic character of $\tilde{A}_c$ we mean a holomorphic function $\xi \neq 0$ on $\tilde{A}_c$ such that $\xi(ab) = \xi(a)\xi(b)$ $(a, b \in \tilde{A}_c)$. ξ being such a character, we denote by $\mathfrak{H}_\xi$ the space of all holomorphic functions f on W such that $f(l_{na}w) = f(w)\xi(a)$ $(n \in N_c^-,\ a \in \tilde{A}_c,\ w \in W)$. Our object now is to prove the following theorem.

THEOREM 1. *$\mathfrak{H}_\xi = \{0\}$ unless there exists a linear function Λ on $\mathfrak{h}$ with the following two properties:* (1) *$\xi(\exp H) = e^{\Lambda(H)}$ $(H \in \mathfrak{h})$ and* (2) *$\Lambda(H_\alpha)$ is a nonnegative integer for every positive root which is not totally positive.*[3]

For any $\bar{x} \in G_0$ and $z \in G_c$ put $z^{\bar{x}} = \bar{x}z\bar{x}^{-1}$. If y is a fixed element in G, xyx^{-1} $(x \in G)$ depends only on $\bar{x}$ and so we may denote it by $y^{\bar{x}}$. Similarly if $w = (s, x) \in S \times G = W$ and $h \in A$, $r_{h^{-1}} l_h w = (s^{\bar{h}}, x^{\bar{h}})$ and so for fixed w it depends only on $\bar{h}$. We shall denote it by $w^{\bar{h}}$. It is clear from its definition that $w \to w^{\bar{h}}$ $(w \in W)$ is a holomorphic mapping of W. A_0 being compact, we normalize its Haar measure $d\bar{h}$ in such a way that $\int_{A_0} d\bar{h} = 1$. We now need a few lemmas.

LEMMA 6. *There exists a function $\psi \in \mathfrak{H}_\xi$ such that*

$$\int_{A_0} \phi(r_x w^{\bar{h}})\, d\bar{h} = \phi(x)\psi(w) \qquad (w \in W, x \in G)$$

for every $\phi \in \mathfrak{H}_\xi$. This function is unique and $\psi(1) = 1$ if $\mathfrak{H}_\xi \neq \{0\}$.

Let us first make some preliminary remarks. Every element X in $\mathfrak{g}$ defines a right-invariant holomorphic infinitesimal transformation [4] X' as

---

[5] For A contains the center of K (Weyl [9]) and K contains the center of G [7].

follows. f being a function which is defined and holomorphic around some point[6] $z \varepsilon G_c$,

$$X'f(z) = \{df(\exp(-tX)z)/dt\}_{t=0}.$$

If X, Y, Z are three elements in $\mathfrak{g}$ and $Z = [X, Y]$, it is easy to check that $Z'f = X'Y'f - Y'X'f$ in some neighborhood of z. Now since $\bar{W}$ is an open submanifold of G_c and since the mapping ν of W onto $\bar{W}$ is everywhere regular, there is exactly one holomorphic infinitesimal transformation on W ν-related to X' [4, Chap. III, § V] which we denote again by X'. Let V be an open set either in W or $\bar{W}$ and let $\mathcal{E}$ be the space of all holomorphic functions on V. Then if we associate to each $X \varepsilon \mathfrak{g}$ the linear mapping $f \to X'f$ $(f \varepsilon \mathcal{E})$ of $\mathcal{E}$ into itself, we get a representation of $\mathfrak{g}$ on $\mathcal{E}$ which can be extended (uniquely) to a representation of the universal enveloping algebra $\mathfrak{B}$ of $\mathfrak{g}$. For any $b \varepsilon \mathfrak{B}$ we denote by b' the corresponding operator on $\mathcal{E}$.

Let w be a point in G_c and f a function which is defined and holomorphic in some neighbourhood of w. Then if $X \varepsilon \mathfrak{g}$ and t is a complex variable, the function $F(t) = f(\exp(-tX)w)$ is defined and holomorphic around the origin in the complex plane and

$$\{(d^m/dt^m)F(t)\}_{t=0} = (X^m)'f(w).$$

Now let $(X_1, \cdots, X_n)$ be a base of $\mathfrak{g}$ over C. We put $X(z) = z_1X_1 + \cdots + z_nX_n$ where $z_1, \cdots, z_n$ are complex numbers. Then the function $F(z) = f((\exp - X(z))w)$ is defined and holomorphic around the origin in C^n. Let $M = (m_1, \cdots, m_n)$ be any sequence of n nonnegative integers. We write $z^M = z_1^{m_1}z_2^{m_2}\cdots z_n^{m_n}$, $M! = m_1!m_2!\cdots m_n!$, $|M| = m_1 + \cdots + m_n$ and

$$\partial^M F/\partial z^M = (\partial^{m_1+\cdots+m_n}/\partial z_1^{m_1}\cdots\partial z_n^{m_n})F.$$

Then if δ is a sufficiently small positive number, $F(z)$ is defined and holomorphic for all (z) such that $|z| = \max_i |z_i| < \delta$ and therefore

$$F(z) = \sum_M F(M)z^M/M! \qquad (|z| < \delta)$$

where $F(M) = (\partial^M F/\partial z^M)_0$, the suffix 0 denoting the value at the origin. Now replace z_i by tz_i where t is a complex number and $|t| \leqq 1$. Then

$$F(tz) = \sum_M F(M)t^{|M|}z^M/M!.$$

[6] Here $X'f(z)$ denotes the value of $X'f$ at z. A similar notation will be used in other cases as well.

On the other hand if $z=(z_1,\cdots,z_n)$ is fixed,

$$F(tz)=\sum_{m\geqq 0}(t^m/m!)(X'(z))^m f(w)$$

for $|t|$ sufficiently small. (Here $X'(z)=(X(z))'$). This follows from the fact mentioned above that

$$\{(d^m/dt^m)F(tz)\}_{t=0}=(X'(z))^m f(w).$$

Hence comparing coefficients of powers of t we get

$$(X'(z))^m f(w)=m!\sum_{m=|M|}F(M)z^M/M!.$$

Since this is true for all sufficiently small values of $|z|$, we can compare the coefficients of z^M on both sides and conclude that $F(M)=X'(M)f(w)$, where $X(M)$ is the coefficient (in $\mathfrak{B}$) of z^M in $(X(z))^m/m!$ $(m=|M|)$ and $X'(M)=(X(M))'$. This proves that

$$f((\exp-X(z))w)=\sum_M X'(M)f(w)z^M/M!$$

for all sufficiently small values of $|z|$.

Now we return to the lemma. Let ϕ and x be fixed elements in $\mathfrak{H}_\xi$ and G respectively. Put $f(w)=\phi(r_x w)$ $(w\,\varepsilon\,W)$. δ being a positive real number, let Q_δ denote the cube in C^n consisting of all points z with $|z|<\delta$. We assume that δ is so small that the following conditions are fulfilled: (1) $\exp(-X(z))\,\varepsilon\,\bar{W}$ for $z\,\varepsilon\,Q_\delta$; (2) the mapping $z\rightarrow\exp(-X(z))$ is regular and therefore open on Q_δ and hence the set $\bar{V}=\exp Q_\delta$ is an open connected neighbourhood of 1 in $\bar{W}$; (3) $\bar{V}$ is evenly covered [4, Chap. II, § VI] under the mapping ν of W onto $\bar{W}$. Let V denote the component of 1 in $\nu^{-1}(\bar{V})$. Define a function $\bar{f}$ on $\bar{V}$ by the rule $\bar{f}(\nu(w))=f(w)$ $(w\,\varepsilon\,V)$. Then f is holomorphic on $\bar{V}$. For any $\bar{h}\,\varepsilon\,A_0$ we extend $Ad(\bar{h})$ to an automorphism of $\mathfrak{B}$ and put $b^{\bar{h}}=Ad(\bar{h})b$ $(b\,\varepsilon\,\mathfrak{B})$. Let $z_i(X)$ $i=1,\cdots,n$ denote the coordinates of $X\,\varepsilon\,\mathfrak{g}$ with respect to the base $(X_1,\cdots,X_n)$. If ϵ is a positive number we denote by $\mathfrak{g}_\epsilon$ the set of all $X\,\varepsilon\,\mathfrak{g}$ such that $|z(X)|=\max_i|z_i(X)|<\epsilon$. Since A_0 is compact, we can choose ϵ so small that if $|z(X)|\leqq\epsilon$, $|z(X^{\bar{h}})|\leqq\delta/2$ for every $\bar{h}\,\varepsilon\,A_0$. Then since $\bar{f}(\exp(-X(z)))$ is holomorphic on Q_δ, it follows from our result above that

$$\bar{f}(\exp(-X))=\sum_{m\geqq 0}(X^m)'\bar{f}(1)/m! \qquad (X\,\varepsilon\,\mathfrak{g}_\delta)$$

and therefore

$$\bar{f}(\exp(-X^{\bar{h}}))=\sum_{m\geqq 0}(1/m!)((X^{\bar{h}})^m)'\bar{f}(1)/m! \qquad (X\,\varepsilon\,\mathfrak{g}_\epsilon,\bar{h}\,\varepsilon\,A_0).$$

For any $z \varepsilon Q_\delta$, let $w(z)$ denote the unique point in V such that $\nu(w(z)) = \exp(-X(z))$. Then if $z(X)$ denotes the point $(z_1(X), \cdots, z_n(X))$ in C^n, it is clear that $w(z(X^{\bar{h}})) = (w(z(X)))^{\bar{h}}$ $(X \varepsilon \mathfrak{g}_\epsilon, \bar{h} \varepsilon A_0)$ and therefore

$$f((w(z))^{\bar{h}}) = \sum_M ((X(M))^{\bar{h}})' f(1) z^M/M! \qquad (|z| < \epsilon, \bar{h} \varepsilon A_0).$$

Moreover in view of our choice of ϵ, for a fixed z the convergence is uniform with respect to $\bar{h}$. Hence

$$\int_{A_0} f((w(z))^{\bar{h}}) d\bar{h} = \sum_M (z^M/M!) \int_{A_0} ((X(M))^{\bar{h}})' f(1) d\bar{h} \qquad (|z| < \epsilon).$$

Now let $q(M) = \int_{A_0} (X(M))^{\bar{h}} d\bar{h}$. Here the integral is well defined since the elements $(X(M))^{\bar{h}}$ span a finite-dimensional subspace of $\mathfrak{B}$. Then

$$\int_{A_0} f((w(z))^{\bar{h}}) d\bar{h} = \sum_M z^M q'(M) f(1)/M! \qquad (|z| < \epsilon).$$

It is obvious from the definition of $q(M)$ that $[H, q(M)] = 0$ if $H \varepsilon \mathfrak{h}$. Hence by the consideration of ranks (see § 7 of [5(f)]) we see that there is exactly one element $h(M)$ in the subalgebra $\mathfrak{H}$ of $\mathfrak{B}$ generated by $(1, \mathfrak{h})$ such that $q(M) \equiv h(M) \bmod \mathfrak{B}\mathfrak{n}_-$. On the other hand $f(l_n w) = f(w)$ if $n \varepsilon N_c^-$ and therefore it follows easily that $b' f(w) = 0$ if $b \epsilon \mathfrak{B}\mathfrak{n}_-$ and $w \varepsilon W$. Moreover since $\tilde{A}_c$ is an abelian Lie group with the Lie algebra $\mathfrak{h}$, there exists a linear function Λ on $\mathfrak{h}$ such that $\xi(\exp H) = e^{\Lambda(H)}$ $(H \varepsilon \mathfrak{h})$. Then $f(l_a w) = e^{\Lambda(H)} f(w)$ if $a = \exp H \varepsilon \tilde{A}_c$. We extend the mapping $H \to -\Lambda(H)$ $(H \varepsilon \mathfrak{h})$ to a (uniquely determined) homomorphism μ_Λ of $\mathfrak{H}$ into C such that $\mu_\Lambda(1) = 1$. Then it is clear that $h' f(w) = \mu_\Lambda(h) f(w)$ for $h \varepsilon \mathfrak{H}$ and $w \varepsilon W$. Hence

$$q'(M) f(1) = h'(M) f(1) = \mu_\Lambda(h(M)) f(1),$$

and therefore

$$\int_{A_0} f((w(z))^{\bar{h}}) d\bar{h} = \sum_M z^M \mu_\Lambda(h(M)) f(1)/m! \qquad (|z| < \epsilon).$$

Now if we put $\Phi(x, w) = \int_{A_0} \phi(r_x w^{\bar{h}}) d\bar{h}$ and recall that $f(1) = \phi(x)$ we get

$$\Phi(x, w(z)) = \phi(x) \sum_M z^M \mu_\Lambda(h(M))/M!$$

provided $|z| < \epsilon$. It is clear that the set of all points $w(z)$ $(|z| < \epsilon)$ is a neighbourhood of 1 in W. Hence from the principle of analytic continuation, there exists at most one holomorphic function ψ on W such that

$$\psi(w(z)) = \sum_M z^M \mu_\Lambda(h(M))/M!$$

if $|z| < \epsilon$. On the other hand if $\mathfrak{H}_\xi \neq \{0\}$ we can choose a function $\phi_0 \neq 0$ in $\mathfrak{H}_\xi$. Since every element in W is of the form $l_s x$ $(s \varepsilon S, x \varepsilon G)$, it is obvious that $\phi_0(x_0) \neq 0$ for some $x_0 \varepsilon G$. Moreover since A_0 is compact, the function

$$\Phi_0(x_0, w) \doteq \int_{A_0} \phi_0(r_{x_0} w^{\bar{h}})\, d\bar{h} \qquad (w \varepsilon W)$$

is obviously holomorphic on W and therefore the same holds for $\Phi_0(x_0, w)/\phi_0(x_0)$. But if we apply the above relation to ϕ_0, we get

$$\Phi_0(x_0, w(z))/\phi_0(x_0) = \sum_M z^M \mu_\Lambda(h(M))/M! \qquad (|z| < \epsilon).$$

This shows that the function ψ certainly exists (if $\mathfrak{H}_\xi \neq \{0\}$). On the other hand it is obvious that $\Phi_0(x_0, l_{na} w) = \xi(a)\Phi_0(x_0, w)$ $(n \varepsilon N_c^-, a \varepsilon \tilde{A}_c, w \varepsilon W)$ and therefore $\psi \varepsilon \mathfrak{H}_\xi$. Furthermore since $\Phi(x, w)$ and $\phi(x)\psi(w)$ coincide on a neighbourhood of 1 in W and since they are both holomorphic in w, they must coincide everywhere. This proves that

$$\int_{A_0} \phi(r_x w^{\bar{h}})\, d\bar{h} = \phi(x)\psi(w) \qquad (x \varepsilon G, w \varepsilon W),$$

and the uniqueness of ψ is obvious from this formula. In particular if we put $\phi = \phi_0$, $x = x_0$ and $w = 1$, we get $\psi(1) = 1$. Finally if $\mathfrak{H}_\xi = \{0\}$, ψ must also be zero and so ψ is unique in any case. Thus the lemma is proved.

Let Λ denote, as above, the linear function on $\mathfrak{h}$ such that $\xi(\exp H) = e^{\Lambda(H)}$ $(H \varepsilon \mathfrak{h})$.

Lemma 7. *Let $\phi \neq 0$ be a function in $\mathfrak{H}_\xi$. Suppose there exists a linear function Λ' on $\mathfrak{h}$ such* that

$$\phi(r_a w) = e^{\Lambda'(H)}\phi(w) \qquad (w \varepsilon W, H \varepsilon \mathfrak{h}_0)$$

where $a = \exp H \varepsilon A$. Then $\Lambda - \Lambda'$ is a linear combination of positive roots with coefficients which are nonnegative integers.

Let V be an open connected neighbourhood of 1 in W such that V is mapped in a one-one fashion on $\bar{V} = \nu(V)$ under the mapping ν. Define a function $\bar{\phi}$ on $\bar{V}$ by setting $\bar{\phi}(\nu(w)) = \phi(w)$ $(w \varepsilon V)$. Let $\alpha_1, \cdots, \alpha_r$ be all the distinct positive roots of $\mathfrak{g}$. We put $X_i = X_{\alpha_i}$ $1 \leqq i \leqq r$. Then $(X_1, \cdots, X_r)$ is a base for $\mathfrak{n}_+$ over C. Put $X(z) = z_1 X_1 + \cdots + z_r X_r$ (where the z_i's are complex numbers) and for any positive ϵ let $\mathfrak{n}_+(\epsilon)$ denote the

neighbourhood of zero in $\mathfrak{n}_+$ consisting of all $X(z)$ with $|z| = \max_i |z_i| < \epsilon$. We can choose ϵ so small that $\exp(-X(z)) \in \bar{V}$ and

$$\bar{\phi}(\exp(-X(z))) = \sum_M X'(M)\bar{\phi}(1)z^M/M!$$

if $|z| < \epsilon$. (Here the notation is similar to what we used in the proof of Lemma 6. M is a sequence $(m_1, \cdots, m_r)$ of r nonnegative integers and $X(M)$ is the coefficient (in $\mathfrak{B}$) of z^M in $(1/m!)(X(z))^m$ where $m = m_1 + \cdots + m_r$. Moreover $M! = m_1! \cdots m_r!$). Choose a positive real δ such that $(X(z))^{\bar{h}} \in \mathfrak{n}_+(\epsilon)$ for all $\bar{h} \in A_0$ if $|z| < \delta$. Then it is obvious that

$$\bar{\phi}(\exp(-X(z))^{\bar{h}}) = \sum_M ((X(M))^{\bar{h}})'\bar{\phi}(1)z^M/M! \qquad (|z| < \delta, \bar{h} \in A_0).$$

However if $h = \exp H$ $(H \in \mathfrak{h}_0)$, it is obvious that

$$\phi(w^{\bar{h}}) = \phi(l_h r_{h^{-1}} w) = \exp(\Lambda(H) - \Lambda'(H))\phi(w) \qquad (w \in W).$$

Hence if $|z| < \delta$,

$$\bar{\phi}(\exp(-X(z))^{\bar{h}}) = \exp(\Lambda(H) - \Lambda'(H))\bar{\phi}(\exp(-X(z))).$$

On the other hand if $M = (m_1, \cdots, m_r)$, it is clear that

$$(X(M))^{\bar{h}} = e^{\alpha_M(H)}X(M),$$

where $\alpha_M = m_1\alpha_1 + \cdots + m_r\alpha_r$. Therefore

$$\exp(\Lambda(H) - \Lambda'(H)) \sum_M X'(M)\bar{\phi}(1)z^M/M! = \sum_M e^{\alpha_M(H)}X'(M)\bar{\phi}(1)z^M/M!$$

for all $H \in \mathfrak{h}_0$ and all $z = (z_1, \cdots, z_r)$ with $|z| < \delta$. Hence comparing coefficients of z^M we get $X'(M)\bar{\phi}(1) = 0$ unless $\Lambda - \Lambda' = \alpha_M$. On the other hand $X'(M)\bar{\phi}(1)$ cannot be zero for all M. For otherwise $\bar{\phi}(\exp(-X(z)) = 0$ if $|z| < \delta$. But since $\mathfrak{g} = \mathfrak{n}_- + \mathfrak{h} + \mathfrak{n}_+$, the elements of the form $na\exp(-X(z))$ $(a \in A_c, n \in N_c^-, |z| < \delta)$ cover a neighbourhood of 1 in $\bar{W}$. Since $\phi(l_{na}w) = \xi(a)\phi(w)$, it is obvious that

$$\bar{\phi}(na\exp(-X(z))) = \xi(a)\bar{\phi}(\exp(-X(z)) = 0$$

if $na\exp(-X(z)) \in \bar{V}$ and $|z| < \delta$. This shows that $\bar{\phi}$ vanishes identically on a neighbourhood of 1 in $\bar{V}$ and therefore ϕ is also zero on some neighbourhood of 1 in W. But then ϕ, being holomorphic, must be zero everywhere on W. Since this contradicts our hypothesis, $\Lambda - \Lambda' = \alpha_M$ for some M and so the lemma is proved.

Every $X \in \mathfrak{g}$ may be regarded as a (left-invariant) holomorphic infini-

tesimal transformation on G_c [4, Chap. IV] and therefore also on its open submanifold $\bar{W}$. Then there is exactly one holomorphic infinitesimal transformation on W which is ν-related to X. We denote it also by X. Let V be an open set either in W or $\bar{W}$ and let $\mathcal{E}$ be the space of all holomorphic functions on V. Then these operations of $\mathfrak{g}$ define a representation of $\mathfrak{g}$ on $\mathcal{E}$ which may be extended uniquely to a representation of $\mathfrak{B}$. If $b \varepsilon \mathfrak{B}$ and $f \varepsilon \mathcal{E}$, we denote by $bf(w)$ the value of bf at w $(w \varepsilon V)$. Let $\mathfrak{X}$ be the subalgebra of $\mathfrak{B}$ generated by $(1, \mathfrak{k})$.

LEMMA 8. *Let ψ be the function of Lemma 6. Then $H\psi = \Lambda(H)\psi$ for $H \varepsilon \mathfrak{h}$ and $X_\alpha\psi = 0$ for every positive root α. Moreover the functions $b\psi$ $(b \varepsilon \mathfrak{X})$ span a finite-dimensional subspace of $\mathfrak{H}_\xi$.*

Since r_u and l_v $(u \varepsilon W_r, v \varepsilon W_l)$ commute, it is an easy matter to verify that if $f \varepsilon \mathfrak{H}_\xi$ and $X \varepsilon \mathfrak{g}$ then Xf is also in $\mathfrak{H}_\xi$. Therefore we get a representation of $\mathfrak{B}$ on $\mathfrak{H}_\xi$. Now for any $\phi \varepsilon \mathfrak{H}_\xi$, consider the function

$$\Phi(x, w) = \phi(r_x w) \qquad (x \varepsilon G, w \varepsilon W)$$

on $G \times W$. Since $(x, w) \rightarrow \nu(r_x w) = \nu(w)\bar{x}$ is a (real) analytic mapping of $G \times W$ into $\bar{W}$, it is clear that $(x, w) \rightarrow r_x w$ is also an analytic mapping of $G \times W$ into W. If $W \varepsilon \mathfrak{g}_0$ it is obvious from the definition of $X\phi$ that

$$X\phi(w) = \{d\Phi(\exp tX, w)/dt\}_{t=0} \qquad (t \varepsilon R).$$

Moreover if $X, Y \varepsilon \mathfrak{g}_0$ and $Z = X + (-1)^{\frac{1}{2}} Y \varepsilon \mathfrak{g}$, $Z\phi = X\phi + (-1)^{\frac{1}{2}}(Y\phi)$. Therefore if Λ' is a linear function on $\mathfrak{h}$, it follows from the above differential equation that $H\phi = \Lambda'(H)\phi$ for every $H \varepsilon \mathfrak{h}$ if and only if

$$\Phi(\exp H, w) = e^{\Lambda'(H)}\phi(w)$$

for all $H \varepsilon \mathfrak{h}_0$ and $w \varepsilon W$. In particular if we apply this criterion to ψ and take into account the fact (which follows from Lemma 6) that $\psi(w^{\bar{h}}) = \psi(w)$ $(\bar{h} \varepsilon A_0, w \varepsilon W)$, we get $H\psi = \Lambda(H)\psi$ $(H \varepsilon \mathfrak{h})$. But then if $\phi = X_\alpha\psi$ it is clear that $H\phi = (\Lambda(H) + \alpha(H))\phi$ $(H \varepsilon \mathfrak{h})$ and therefore by the above criterion

$$\phi(r_h w) = \exp(\Lambda(H) + \alpha(H))\phi(w) \qquad (H \varepsilon \mathfrak{h}_0, w \varepsilon W)$$

where $h = \exp H \varepsilon A$. α being a positive root, we can conclude from Lemma 7 that $\phi = X_\alpha\psi = 0$.

To prove the last assertion we may assume that $\mathfrak{H}_\xi \neq \{0\}$. Let K and K' be the analytic subgroups of G corresponding to $\mathfrak{k}_0$ and $\mathfrak{k}_0' = [\mathfrak{k}_0, \mathfrak{k}_0]$ respectively. Then K' is compact and semisimple [4] and the function

$$\Psi(u, w) = \psi(r_u w) \qquad (u \varepsilon K', w \varepsilon W)$$

is continuous on $K' \times W$. Moreover $\Psi(1,1) = \psi(1) = 1$. Therefore from the Peter-Weyl Theorem for K', there exists an irreducible character χ of K' with the property that the function

$$\Psi'(u,w) = \int_{K'} \chi(v^{-1}) \Psi(uv,w)\, dv$$

is not identically zero on $K' \times W$. (Here dv is the Haar measure on K'). Since $\Psi(uv,w) = \Psi(v, r_u w)$, it follows that $\Psi'(u,w) = \Psi'(1, r_u w)$. Therefore if $\psi'(w) = \Psi'(1,w)$ $(w \in W)$, it is obvious that $\psi' \in \mathfrak{H}_\xi$ and it is not zero. Now if $h \in A$,

$$\psi'(w^{\bar{h}}) = \int_{K'} \chi(v^{-1})\, \psi(r_v w^{\bar{h}})\, dv.$$

Moreover K' is a normal subgroup of K and it is clear that $\chi(h^{-1}v^{-1}h) = \chi(v^{-1})$. Therefore

$$\int_{K'} \chi(v^{-1}) \psi(r_v w^{\bar{h}})\, dv = \int_{K'} \chi(v^{-1}) \psi((r_v w)^{\bar{h}})\, dv = \int_{K'} \chi(v^{-1}) \psi(r_v w)\, dv = \psi'(w),$$

since $\psi(z^{\bar{h}}) = \psi(z)$ $(z \in W)$. This shows that $\psi'(w^{\bar{h}}) = \psi'(w)$ and therefore from Lemma 6,

$$\psi'(w) = \int_{A_0} \psi'(w^{\bar{h}})\, d\bar{h} = \psi'(1)\psi(w) \qquad (w \in W).$$

This proves that the function ψ and ψ' differ only by a constant factor which however cannot be zero since neither ψ nor ψ' is zero. For any fixed $u \in K'$ put $\psi_u'(w) = \Psi'(u,w) = \psi'(r_u w)$. Then it is clear from the definition of $\Psi'(u,w)$ that $\psi_u' \in \mathfrak{H}_\xi$ and the dimension of the subspace V of $\mathfrak{H}_\xi$ spanned by all ψ_u' $(u \in K')$ is finite. For any $\phi \in V$ define $\phi_u(w) = \phi(r_u w)$ $(u \in K', w \in W)$. Then if to each $u \in K'$ we associate the linear mapping $\sigma(u) : \phi \to \phi_u$ of V into itself, we get a representation σ of K' on the finite-dimensional space V. We denote the corresponding representation of $\mathfrak{k}_0'$ also by σ. Since $\sigma(u)\phi(w) = \phi(r_u w)$ it follows immediately that $\sigma(X)\phi = X\phi$ $(X \in \mathfrak{k}_0', \phi \in V)$. On the other hand if $\mathfrak{c}_0$ is the center of $\mathfrak{k}_0$ and $h = \exp H \in A$ $(H \in \mathfrak{c}_0)$ it is clear that $\psi_u(r_h w) = \psi(r_h r_u w) = e^{\Lambda(H)}\psi(r_u w) = e^{\Lambda(H)}\psi_u(w)$. Since V is spanned by the functions ψ_u, we can conclude that $\phi(r_h w) = e^{\Lambda(H)}\phi(w)$ for all $\phi \in V$. But as we have seen earlier this implies that $H\phi = \Lambda(H)\phi$ and therefore V is invariant under the operations of $\mathfrak{k}_0' + \mathfrak{c}_0 = \mathfrak{k}_0$ and so also under those of $\mathfrak{k}$. Since $\psi \in V$, the last assertion of the Lemma follows immediately.

It is now obvious that Theorem 1 is a direct consequence of Lemma 8 and Theorem 1 of [5(f)].

5. Representations on a Hilbert space of holomorphic functions. We shall now prove a converse of Theorem 1. Since $\tilde{A}_c$ is the direct product of A_+ and A and A_+ is simply connected, $\exp H = 1$ in $\tilde{A}_c$ $(H \varepsilon \mathfrak{h})$ if and only if $H \varepsilon \mathfrak{h}_0$ and $\exp H = 1$ in A. On the other hand if D and A' are the analytic subgroups of A corresponding to[3] $\mathfrak{c}_0$ and $\mathfrak{h}_0' = \mathfrak{h}_0 \cap \mathfrak{k}_0'$, A is the direct product of D and A' and D is simply connected.[4] Therefore $\exp H \neq 1$ in A unless $H \varepsilon \mathfrak{h}_0'$. Now suppose $H \varepsilon \mathfrak{h}_0'$. Since $A' \subset K'$ and K' is compact, $\exp H = 1$ if and only if it lies in the kernel of every finite-dimensional irreducible representation of K'. But K' is simply connected[4] and therefore there is a one-one correspondence between finite-dimensional representations of K' and those of $\mathfrak{k}'$. So it is clear that $\exp H = 1$ if and only if $e^{\Lambda(H)} = 1$ for all weights Λ (with respect to $\mathfrak{h}'$) of all finite-dimensional representations of $\mathfrak{k}'$. Since we may identify compact roots of $\mathfrak{g}$ with roots of $\mathfrak{k}'$, it follows[7] that a linear function Λ on $\mathfrak{h}$ coincides on $\mathfrak{h}'$ with the weight of some finite-dimensional representation of $\mathfrak{k}'$, if and only if $\Lambda(H_\alpha)$ is an integer for every compact root α. Therefore *in order that there should exist a (holomorphic) character ξ of $\tilde{A}_c$ satisfying the equation $\xi(\exp H) = e^{\Lambda(H)}$ $(H \varepsilon \mathfrak{h})$ it is necessary and sufficient that $\Lambda(H_\alpha)$ should be an integer for every compact root α.*

Now let μ and ω be two nonnegative (Haar) measurable functions on G_0. We assume that μ is not identically zero, ω is bounded on every compact set and $\mu(\bar{x}\bar{y}) \leqq \mu(\bar{x})\omega(\bar{y})$ $(\bar{x}, \bar{y} \varepsilon G_0)$. Then $\mu(\bar{y}) \leqq \mu(\bar{x})\omega(\bar{x}^{-1}\bar{y})$ and since μ is not identically zero it follows that $\mu(\bar{x}) \neq 0$. Moreover

$$\{\mu(\bar{x}\bar{y})\}^{-1} \leqq \{\mu(\bar{x})\}^{-1}\omega(\bar{y}^{-1})$$

and since $\omega(\bar{y})$ (and therefore also $\omega(\bar{y}^{-1})$) is bounded on every compact set, it follows from the above inequalities that the same holds for both μ and $1/\mu$. Hence if $d\bar{x}$ is the Haar measure on G_0, the two measures $\mu(\bar{x})d\bar{x}$ and $d\bar{x}$ are absolutely continuous with respect to each other.

Let Λ be a *real*[2] linear function on $\mathfrak{h}$. We assume that $\Lambda(H_\alpha)$ is a nonnegative integer for every positive root α which is not totally positive. Then, as we have seen above, there exists a holomorphic character ξ of $\tilde{A}_c$ such that $\xi(\exp H) = e^{\Lambda(H)}$ $(H \varepsilon \mathfrak{h})$. Let f be a complex-valued function on W such that $f(l_a w) = \xi(a)f(w)$ for all $a \varepsilon \tilde{A}_c$ and $w \varepsilon W$. If Z is the center of G, $Z \subset A$[5] and therefore $f(l_z w) = \xi(z)f(w)$ $(z \varepsilon Z)$ and $|\xi(z)| = 1$ because Λ is real. This shows that $|f(zx)| = |f(x)|$ $(z \varepsilon Z, x \varepsilon G)$ and so $|f(x)|$ depends only on $\bar{x}$. Hence if $|f(x)|$ happens to be a measurable function of $\bar{x}$,

[7] This can be deduced easily from Theorem 1 of [5(a)]. See also Weyl [9].

we can consider the integral $\int_{G_0} |f(x)|^2 \mu(\bar{x}) d\bar{x}$. Now let $\mathfrak{H}_\Lambda(\mu)$ denote the space of all holomorphic functions f on W satisfying the following two conditions:

(1) $\quad f(l_{na}w) = \xi(a) f(w) \qquad (n \in N_c^-, a \in \tilde{A}_c, w \in W)$

(2) $\quad \| f \|^2 = \int_{G_0} |f(x)|^2 \mu(\bar{x}) d\bar{x} < \infty.$

We put $\mathfrak{H} = \mathfrak{H}_\Lambda(\mu)$ for convenience. Then every $f \in \mathfrak{H}$ is certainly continuous on G. Therefore since the measures $d\bar{x}$ and $\mu(\bar{x}) d\bar{x}$ are absolutely continuous with respect to each other, $\| f \|$ is positive unless f vanishes identically on G. But then in view of condition (1) above, $f = 0$. Therefore in order to prove that $\mathfrak{H}$ is a Hilbert space, it only remains to prove that it is complete. After this has been done we intend to define a representation of G on $\mathfrak{H}$ and prove that $\mathfrak{H} = \mathfrak{H}_\Lambda(\mu) \neq \{0\}$ for a suitable choice of μ.

Let x_0 be a fixed element in G and let $X_1, \cdots, X_n$ be a base for $\mathfrak{g}$ over C. If $X \in \mathfrak{g}$, we denote by $z_1(X), \cdots, z_n(X)$ the coordinates of X with respect to this base and for any positive number ϵ we define (as before) $\mathfrak{g}_\epsilon$ to be the set of all $X \in \mathfrak{g}$ such that $|z(X)| = \max_i |z_i(X)| < \epsilon$. Now choose ϵ so small that the following conditions hold: (1) the mapping $X \to \exp(-X)$ of $\mathfrak{g}$ into G_c is univalent and regular on $\mathfrak{g}_\epsilon$, (2) $\exp(-X)\bar{x}_0 \in \bar{W}$ for $X \in \mathfrak{g}_\epsilon$, (3) the set $\bar{V} = \exp(\mathfrak{g}_\epsilon)\bar{x}_0$ is evenly covered under the mapping ν of W onto $\bar{W}$. Let V denote the connected component of x_0 in $\nu^{-1}(\bar{V})$. Then V and $\bar{V}$ are both open and ν defines a one-one regular holomorphic mapping of V onto $\bar{V}$.

For any $X \in \mathfrak{g}$, consider the endomorphism

$$(1 - \exp(-adX))/adX = \sum_{m \geqq 0} (-1)^m (adX)^m/(m+1)!$$

of the complex vector space $\mathfrak{g}$. We denote by $\Delta(X)$ its determinant. Then $\Delta(X)$ is clearly a holomorphic function on $\mathfrak{g}$ and a well-known computation [4, p. 157] shows that if $x^* = \exp(-X)\bar{x}_0$,

$$dx^* = d(x^*)^{-1} = |\Delta(X)|^2 |dX|^2 \qquad (X \in \mathfrak{g}_\epsilon)$$

where dx^* is the Haar measure on G_c and $|dX|^2$ is the Euclidean measure on $\mathfrak{g}$ (regarded as a vector space over R). Since $\Delta(0) = 1$, we may assume that ϵ is so small that $|\Delta(X)| \geqq \frac{1}{2}$ on $\mathfrak{g}_\epsilon$.

On the other hand we know (see Corollary to Lemma 1 and Lemma 26 of [5(b)]) that $(n, a, \bar{x}) \to na\bar{x}$ $(n \in N_c^-, a \in A_+, \bar{x} \in G_0)$ is a one-one regular mapping of $N_c^- \times A_+ \times G_0$ onto $\bar{W}$ and an easy computation shows that

$$dx^* = e^{4\rho(\log a)} dn da d\bar{x},$$

where $x^* = na\bar{x}$ and dx^*, dn, da are the (suitably normalised) Haar measures of G_c, N_c^- and A_+ respectively. Let V' be the set of all points $(n, a, \bar{x}) \varepsilon N_c^- \times A_+ \times G_0$ such that $na\bar{x} \varepsilon \bar{V}$ and let V_1, V_2, V_3 denote the projections of V' on each of the factors N_c^-, A_+, G_0 respectively. Then they are all open and since the closure of $\mathfrak{g}_\epsilon$ is compact, it is clear that the closures of $\bar{V}, V', V_1, V_2, V_3$ are also all compact. We can choose open neighbourhoods V_1', V_2', V_3' of $1, 1, \bar{x}_0$ in N_c^-, A_+, G_0 respectively such that $V_1'V_2'V_3' \subset \bar{V}$. Also choose a positive $\delta < \epsilon$ such that $\exp(-X)\bar{x}_0 \varepsilon V_1'V_2'V_3'$ if $X \varepsilon \mathfrak{g}_\delta$. If f is any function in $\mathfrak{H}$, we denote by $\bar{f}$ the function on $\bar{V}$ given by $\bar{f}(\nu(w)) = f(w)$ $(w \varepsilon V)$. Then

$$\int_{\mathfrak{g}_\delta} |\bar{f}(\exp(-X)\bar{x}_0)|^2 \, |dX|^2 \leqq 4 \int_{\mathfrak{g}_\delta} |\bar{f}(\exp(-X)x_0)|^2 \, |\Delta(X)|^2 \, |dX|^2$$

$$\leqq 4 \int_{V_1 \times V_2 \times V_3} |\bar{f}(na\bar{x})|^2 \, e^{4\rho(\log a)} dn da d\bar{x} = M_1 \int_{V_3} |\bar{f}(\bar{x})|^2 \, d\bar{x}$$

where $M_1 = 4 \int_{V_1} \exp(2\Lambda(\log a) + 4\rho(\log a)) \, da \int_{V_2} dn$. On the other hand since the closure of V_3 is compact, $1/\mu$ is bounded on V_3. Hence

$$\int_{V_3} |\bar{f}(\bar{x})|^2 \, d\bar{x} \leqq M_2 \int_{V_3} |\bar{f}(\bar{x})|^2 \, \mu(\bar{x}) \, d\bar{x} \leqq M_2 \, \| f \|^2$$

where M_2 is an upper bound for μ^{-1} on V_3. This proves that

$$\int_{\mathfrak{g}_\delta} |\bar{f}(\exp(-X)\bar{x}_0)|^2 \, |dX|^2 \leqq M \, \| f \|^2$$

where $M = M_1 M_2$. Since $\bar{f}(\exp(-X)\bar{x}_0)$ is a holomorphic function of X on $\mathfrak{g}_\epsilon$, it follows from a classical argument (see Bochner and Martin [2, p. 117]) that if $\| f \|$ tends to zero, $\bar{f}(\exp(-X)\bar{x}_0)$ tends to zero uniformly on every compact subset of $\mathfrak{g}_\delta$. This proves that $\| f \| \to 0$ implies the uniform convergence of f to zero on some neighbourhood of x_0 in W. Moreover since x_0 can be any point in G and since $f(nax) = \xi(a) f(x)$ $(n \varepsilon N_c^-, a \varepsilon A_+, x \varepsilon G)$ it is clear that the same conclusion holds in some neighbourhood of any given point in W. Therefore if f_m is a Cauchy sequence in $\mathfrak{H}$, it converges uniformly on every compact set in W and hence the limit function f is holomorphic on W and clearly $f(l_{na}w) = \xi(a) f(w)$. Moreover it then follows by well-known elementary arguments that

$$\| f \|^2 = \int_{G_0} |f(x)|^2 \, \mu(\bar{x}) \, d\bar{x} = \lim_{m \to \infty} \| f_m \|^2 < \infty$$

and $\| f - f_m \| \to 0$. This proves that f lies in $\mathfrak{H}$ and f_m converges to f in $\mathfrak{H}$. Therefore $\mathfrak{H}$ is complete.

Now we define a representation π of G on $\mathfrak{H}$ as follows. If $f \varepsilon \mathfrak{H}$ and $y \varepsilon G$, $\pi(y)f$ is the function whose value at w is $f(r_y w)$ $(w \varepsilon W)$. Since $\mu(\bar{x}\bar{y}^{-1}) \leqq \mu(\bar{x})\omega(\bar{y}^{-1})$,

$$\int_{G_0} |f(xy)|^2 \mu(\bar{x})\, d\bar{x} \leqq \omega(\bar{y}^{-1}) \| f \|^2,$$

and so it follows easily that $\pi(y)f \varepsilon \mathfrak{H}$ and $\| \pi(y)f \|^2 \leqq \omega(\bar{y}^{-1}) \| f \|^2$. This shows moreover that the operators $\pi(y)$ remain uniformly bounded on any compact set. Now let $V = V^{-1}$ be any compact neighbourhood of 1 in G and let $\bar{V}$ be its image in G_0 under the mapping $x \to \bar{x}$. For any given $\epsilon > 0$, we can choose a compact set G_1 in G_0 such that

$$\int_{{}^cG_1} |f(x)|^2 \mu(\bar{x})\, d\bar{x} \leqq \epsilon^2.$$

(${}^cG_1$ is the complement of G_1 in G_0). Put $G_2 = G_1\bar{V}$. Then G_2 is also compact and if $y \varepsilon V$, $({}^cG_2)\bar{y} \subset {}^cG_1$. Hence

$$\int_{{}^cG_2} |f(xy)|^2 \mu(\bar{x}) d\bar{x} \leqq \omega(\bar{y}^{-1}) \int_{({}^cG_1)} |f(x)|^2 \mu(\bar{x}) d\bar{x} \leqq M^2\epsilon^2,$$

where M^2 is an upper bound for $\omega(\bar{y})$ on $\bar{V}$. Moreover

$$\| \pi(y)f - f \|^2 = \int_{G_2} |f(xy) - f(x)|^2 \mu(\bar{x}) d\bar{x} + \int_{{}^cG_2} |f(xy) - f(x)|^2 \mu(\bar{x}) d\bar{x}.$$

But

$$\int_{{}^cG_2} |f(xy) - f(x)|^2 \mu(\bar{x}) d\bar{x}$$
$$\leqq [(\int_{{}^cG_2} |f(xy)|^2 \mu(\bar{x}) d\bar{x})^{\frac{1}{2}} + (\int_{{}^cG_2} |f(x)|^2 \mu(\bar{x}) d\bar{x})^{\frac{1}{2}}]^2 \leqq \{(M+1)\epsilon\}^2,$$

and since G_2 is compact,

$$\operatorname*{Lim}_{y \to 1} \int_{G_2} |f(xy) - f(x)|^2 \mu(\bar{x}) d\bar{x} = 0.$$

As ϵ is arbitrary, this shows that $\operatorname*{Lim}_{y \to 1} \| \pi(y)f - f \| = 0$. Moreover it is obvious that $\pi(xy) = \pi(x)\pi(y)$ $(x, y \varepsilon G)$ and therefore π is a representation [5(b), p. 201] of G on $\mathfrak{H}$.

It is obvious that $\mathfrak{H}$ is contained in the space $\mathfrak{H}_\xi$ of Lemma 6.

LEMMA 9. *If* $\mathfrak{H} = \mathfrak{H}_\Lambda(\mu) \neq \{0\}$, *the function* ψ *of Lemma* 6 *lies in* $\mathfrak{H}$.

For if we choose $\phi_0 \neq 0$ in $\mathfrak{H}$, it is clear that $\phi_0(x_0) \neq 0$ for some $x_0 \varepsilon G_0$. Then it follows from Lemma 6 that

$$\int_{A_0} \phi_0(y^{\bar{h}} x_0)\, d\bar{h} = \phi_0(x_0)\psi(y),$$

and therefore

$$|\phi_0(x_0)|^2 \int_{G_0} |\psi(y)|^2 \mu(\bar{y})\, d\bar{y} \leqq \int_{G_0} \mu(\bar{y})\, d\bar{y} \int_{A_0} |\phi_0(y^{\bar{h}} x_0)|^2\, d\bar{h}$$
$$= \int_{A_0} d\bar{h} \int_{G_0} |\phi_0(y^{\bar{h}} x_0)|^2 \mu(\bar{y})\, d\bar{y}.$$

But if $h \varepsilon A$,

$$\int_{G_0} |\phi_0(y^{\bar{h}} x_0)|^2 \mu(\bar{y})\, d\bar{y} = \int_{G_0} |\phi_0(hyh^{-1}x_0)|^2 \mu(\bar{y})\, d\bar{y}$$
$$= \int_{G_0} |\phi_0(y)|^2 \mu(\bar{y}\bar{x}_0^{-1}\bar{h})\, d\bar{y} \leqq \omega(\bar{x}_0^{-1}\bar{h}) \| \phi_0 \|^2,$$

since $|\xi(h)| = 1$. Therefore if M is an upper bound for ω on the compact set $\bar{x}_0^{-1}A_0$, it follows that $|\phi_0(x_0)|^2 \| \psi \|^2 \leqq M \| \phi_0 \|^2$. Since $\phi_0(x_0) \neq 0$, we conclude that $\| \psi \|^2 < \infty$ and therefore $\psi \varepsilon \mathfrak{H}$.

LEMMA 10. *Let ϕ be any element in $\mathfrak{H}$ which is well-behaved*[8] *under π. Then $\pi(X)\phi = X\phi$ $(X \varepsilon \mathfrak{g})$. Moreover if $\mathfrak{H} \neq \{0\}$, ψ is well-behaved under π.*

We have seen that if ϕ tends to ϕ_0 in $\mathfrak{H}$ then $\phi(w)$ tends to $\phi_0(w)$ uniformly on every compact set in W. In particular if ϕ is well-behaved under π and $X \varepsilon \mathfrak{g}_0$, $(1/t)\{\pi(\exp tX)\phi - \phi\}$ $(t \varepsilon R)$ tends to $\pi(X)\phi$ in $\mathfrak{H}$ as $t \rightarrow 0$. On the other hand it is obvious that $(1/t)\{\pi(\exp tX)\phi - \phi\}$ tends to $X\phi$ uniformly on every compact set in W. Therefore $X\phi = \pi(X)\phi$ and by linearity this remains true if $X \varepsilon \mathfrak{g}$. Also we know that $\pi(z) = \xi(z)\pi(1)$ $(z \varepsilon Z)$. Therefore it follows from Theorem 4 of [5(b), p. 224] that the space of well-behaved elements is dense in $\mathfrak{H}$. On the other hand if $h \varepsilon A$, it is obvious that $\xi(h^{-1})\pi(h)$ depends only on $\bar{h}$ and so we may denote it by $\bar{\pi}(\bar{h})$. Then $\bar{h} \rightarrow \bar{\pi}(\bar{h})$ $(\bar{h} \varepsilon A_0)$ is a representation of A_0 on $\mathfrak{H}$ and

$$E = \int_{A_0} \bar{\pi}(\bar{h})\, d\bar{h}$$

is a bounded operator with $E^2 = E$. Moreover since A_0 is compact, it follows easily (see Lemma 29 of [5(b)]) that if ϕ is well-behaved under π, the same is true of $E\phi$. Moreover if we regard

$$E\phi = \int_{A_0} \bar{\pi}(\bar{h}) \phi\, d\bar{h} \qquad (\phi \varepsilon \mathfrak{H})$$

as the limit of a sum in $\mathfrak{H}$, it is clear from the remarks on convergence made above that[6]

$$E\phi(w) = \int_{A_0} \bar{\pi}(\bar{h}) \phi(w)\, d\bar{h} \qquad (w \varepsilon W).$$

[8] We use here and in the rest of this paper the terminology of [5(b)].

But if $h \in A$,

$$\bar{\pi}(\bar{h})\phi(w) = \xi(h^{-1})\pi(h)\phi(w) = \xi(h^{-1})\phi(r_h w) = \xi(h^{-1})\phi(l_h w^{\bar{a}}) = \phi(w^{\bar{a}})$$

where $a = h^{-1}$. Therefore, from Lemma 6,

$$E\phi(w) = \phi(1)\psi(w) \quad \text{or} \quad E\phi = \phi(1)\psi.$$

Now assume that $\mathfrak{H} \neq \{0\}$. Then from Lemma 9, $\psi \in \mathfrak{H}$ and so we can choose a sequence ϕ_m of well-behaved elements in $\mathfrak{H}$ such that $\phi_m \to \psi$ in $\mathfrak{H}$. Since E is bounded

$$\phi_m(1)\psi = E\phi_m \to E\psi = \psi(1)\psi = \psi$$

from Lemma 6. Therefore $\phi_m(1) \neq 0$ if m is sufficiently large and since $E\phi_m$ is well-behaved the same holds for $\psi = \{\phi_m(1)\}^{-1}E\phi_m$.

LEMMA 11. *Suppose $\mathfrak{H} \neq \{0\}$ and $\mathfrak{H}_1$ is the smallest closed subspace of $\mathfrak{H}$ containing ψ which is invariant under $\pi(G)$. Then the representation of G defined on $\mathfrak{H}_1$ under π is irreducible and quasi-simple.*[8]

Let $\mathfrak{H}_2 \neq \{0\}$ be any closed invariant subspace of $\mathfrak{H}_1$. Choose $\phi \neq 0$ in $\mathfrak{H}_2$. Then as we have seen above $E\pi(x)\phi = \{\pi(x)\phi(1)\}\psi = \phi(x)\psi$ $(x \in G)$. Since $\phi \neq 0$ it is clear that $\phi(x) \neq 0$ for some x and therefore $\psi \in \mathfrak{H}_2$. But this implies that $\mathfrak{H}_2 = \mathfrak{H}_1$ and therefore $\mathfrak{H}_2$ is irreducible.

Now let z be any element in $\mathfrak{B}$ of rank[3] zero. Then it is clear that if ϕ is a well-behaved element in $\mathfrak{H}$,

$$\pi(h)\pi(z)\phi = \pi(z)\pi(h)\phi \qquad (h \in A),$$

and therefore $E\pi(z)\phi = \pi(z)E\phi$. In particular

$$E\pi(z)\psi = \pi(z)E\psi = \pi(z)\psi.$$

But we know that $E\phi = \phi(1)\psi$ for every $\phi \in \mathfrak{H}$. Therefore

$$\pi(z)\psi = E\pi(z)\psi = \chi(z)\psi,$$

where $\chi(z)$ is the value of $\pi(z)\psi$ at 1. Now if z lies in the center of $\mathfrak{B}$, it follows that

$$\pi(z)\pi(x)\psi = \pi(x)\pi(z)\psi = \chi(z)\pi(x)\psi \qquad (x \in G).$$

Hence if V is the subspace of $\mathfrak{H}_1$ spanned by $\pi(x)\psi$ $(x \in G)$, $\pi(z)\phi = \chi(z)\phi$ for all $\phi \in V$. Since V consists of well-behaved elements and since V is

dense in $\mathfrak{H}_1$ the quasi-simplicity of the representation on $\mathfrak{H}_1$ now follows from[9] Lemma 32 of [5(b)].

LEMMA 12. *Suppose $\mathfrak{H} \neq \{0\}$ and π is a unitary representation. Then $\mathfrak{H}$ is irreducible under π.*

In view of Lemma 11 it is enough to prove that $\mathfrak{H}_1 = \mathfrak{H}$. Let $\mathfrak{H}_2$ be the orthogonal complement of $\mathfrak{H}_1$ in $\mathfrak{H}$. Since π is unitary, $\mathfrak{H}_2$ is invariant under π. Let ϕ be any element in $\mathfrak{H}_2$. Then we have seen that $E\pi(x)\phi = \phi(x)\psi$ $(x \in G)$. Therefore $\phi(x)\psi \in \mathfrak{H}_2 \cap \mathfrak{H}_1 = \{0\}$ and so $\phi(x) = 0$. This being true for every $x \in G$, it is clear that $\phi = 0$. This proves that $\mathfrak{H}_2 = \{0\}$ and therefore $\mathfrak{H}_1 = \mathfrak{H}$.

6. Computation of the function ψ. We shall now determine the function ψ explicitly. We know from Lemma 4 that $Q = P_c^- M_c P_c^+$ is an open submanifold of G_c and there exists a (unique) holomorphic mapping m of Q into M_c such that $q \in P_c^- m(q) P_c^+$ for every $q \in Q$. Let $\tilde{M}_c$ be the simply connected covering group of M_c and γ the natural homomorphism of $\tilde{M}_c$ onto M_c. Since $\bar{W} \subset Q$, $w \rightarrow m(\nu(w))$ is a holomorphic mapping of W into M_c. But W is simply connected and so there exists a (unique) holomorphic mapping $\tilde{m}$ of W into $\tilde{M}_c$ such that $\tilde{m}(1) = 1$ and $\gamma \circ \tilde{m} = m \circ \nu$. Now if we use the notation of [5(f), § 5] it is clear that $\mathfrak{m} = \mathfrak{g}' + \mathfrak{k}$. Hence if $\mathfrak{c}$ is the center of $\mathfrak{k}$, $\mathfrak{c}_+ = \mathfrak{c} \cap \mathfrak{g}_+$ is the center of $\mathfrak{m}$ and $\mathfrak{m}$ is the direct sum of $\mathfrak{c}_+$ and $\mathfrak{m}' = [\mathfrak{m}, \mathfrak{m}]$. Therefore $\tilde{M}_c$ being simply connected, there exists a (unique) holomorphic mapping Γ of W into $\mathfrak{c}_+$ such that $\tilde{m}(w) \in \tilde{M}_c' \exp \Gamma(w)$. Here $\tilde{M}_c'$ is the analytic subgroup of $\tilde{M}_c$ corresponding to $\mathfrak{m}'$.

LEMMA 13. *Let w be any element in W. Then*

$$\Gamma(l_{na}w) = \Gamma(a) + \Gamma(w),\ \Gamma(r_u w) = \Gamma(w) + \Gamma(u),\ \Gamma(ux) = \Gamma(u) + \Gamma(x)$$

---

[9] It was pointed out to me by Dixmier that the proof of Lemma 33 of [5(b)] is incorrect. Since $\tilde{M}$ is dense in $\bar{\mathfrak{H}}$ only in the weak topology, one cannot conclude that $|(\bar{\phi}, A\psi)| \leqq |B|_N |\bar{\phi}| |\psi|$ for all $\bar{\phi} \in \bar{\mathfrak{H}}$ and $\psi \in M$, knowing it to be true for $\bar{\phi} \in \tilde{M}$ and $\psi \in M$. However suppose there exists a positive real number a such that every element ψ_0 in $\bar{\mathfrak{H}}$ can be approximated arbitrarily well (in the weak topology) by elements $\psi \in \bar{M}$ with $|\psi| \leqq a |\psi_0|$. Then it follows immediately that $|(\bar{\phi}, A\psi)| \leqq a |B|_N |\bar{\phi}| |\psi|$ for all $\bar{\phi} \in \bar{\mathfrak{H}}$ and $\psi \in M$ and therefore $|A|_M \leqq a |B|_N$. Hence, in particular, A is bounded if B is bounded. On the other hand if we use the notation of [5(b), pp. 226-227], it is clear that if n is sufficiently large $|A_{f_n}\psi| \leqq a |\psi|$ where $a = 2 \sup_{x \in \omega} |\pi(x)|$ and ω is a compact set outside which all f_n are zero. From this it follows that $|\bar{A}_{f_n}| \leqq a$ and therefore the subspace $\bar{V}$ of $\bar{\mathfrak{H}}$ does have the above additional property. The proof of Lemma 32 of [5(b)] now goes through without any further modification.

for $n \varepsilon N_c^-$, $a \varepsilon \tilde{A}_c$, $u \varepsilon K$ *and* $x \varepsilon G$. *Moreover if* $a = \exp H \varepsilon \tilde{A}_c$ $(H \varepsilon \mathfrak{H})$ *and* $H = H_+ + H'$ $(H_+ \varepsilon \mathfrak{c}_+,\ H' \varepsilon \mathfrak{h} \cap \mathfrak{m}')$ *then* $\Gamma(a) = H_+$.

Since $[\mathfrak{m}, \mathfrak{p}_-] \subset \mathfrak{p}_-$ and $[\mathfrak{m}, \mathfrak{p}_+] \subset \mathfrak{p}_+$, it follows that $M_c Q = Q M_c = Q$ and $m(vq) = vm(q)$, $m(qv) = m(q)v$ $(v \varepsilon M_c, q \varepsilon Q)$. Moreover $\mathfrak{n}_- \subset \mathfrak{p}_- + \mathfrak{k}' \subset \mathfrak{p}_- + \mathfrak{m}'$. Therefore $N_c^- \subset P_c^- \gamma(\tilde{M}_c')$. Similarly if $u \varepsilon K$, $m(q\bar{u}) = m(q)\bar{u}$ $(q \varepsilon Q)$ and $m(\bar{u}x) = \bar{u}m(x)$ $(x \varepsilon G)$. The first part of the lemma follows immediately from these facts. Now put $a_+ = \exp H_+$ and $a' = \exp H'$. Then $a = a_+ a'$ and therefore $\Gamma(a) = \Gamma(a_+) + \Gamma(a')$. But since $a' \varepsilon \tilde{M}_c'$, $\Gamma(a') = 0$. On the other hand it is obvious that $\Gamma(a_+) = H_+$ and so the lemma is proved.

By a complex representation of G_c we mean a finite-dimensional representation such that the corresponding representation of $\mathfrak{g}$ is linear over C. Since G_c is simply connected we may identify the finite-dimensional representations of $\mathfrak{g}$ with the complex representations of G_c. If V is the representation space of such a representation π, it would be convenient to regard V as a Hilbert space in such a way that π becomes unitary on the subgroup U corresponding to $\mathfrak{u} = \mathfrak{k}_0 + (-1)^{\frac{1}{2}}\mathfrak{p}_0$. Since U is compact, this is always possible. Whenever we speak of a finite-dimensional representation of $\mathfrak{g}$ (or of a complex representation of G_c) we shall tacitly assume that such a Hilbert space structure has already been introduced in the representation space.

Now let Λ be the linear function of Section 5 and let $\alpha_1, \cdots, \alpha_l$ be a fundamental system of positive roots of $\mathfrak{g}$ (see Corollary 2 to Lemma 4 of [5(f)]). We assume that $\alpha_1, \cdots, \alpha_t$ are all the totally positive roots among these. Define a linear function Λ_0 on $\mathfrak{h}$ by the conditions $\Lambda_0(H_{\alpha_i}) = 0$ $1 \leqq i \leqq t$ and $\Lambda_0(H_{\alpha_i}) = \Lambda(H_{\alpha_i})$ $(t < i \leqq l)$. Then Λ_0 is a dominant integral function and therefore from Theorem 1 of [5(a)] there exists an irreducible representation σ of $\mathfrak{g}$ on the finite-dimensional space V with the highest weight Λ_0. Let ϕ_0 be a unit vector in V belonging to the weight Λ_0 and put $\lambda = \Lambda - \Lambda_0$.

LEMMA 14. *Let ξ be the holomorphic character of $\tilde{A}_c$ corresponding to Λ. Then the function ψ of Lemma 6 is given by the formula*

$$\psi(w) = (\phi_0, \sigma(\nu(w))\phi_0)e^{\lambda(\Gamma(w))} \qquad (w \varepsilon W).$$

First observe that since σ is unitary on U, the adjoint of the operator $\sigma(X)$ is $-\sigma(\bar{\theta}(X))$ $(X \varepsilon \mathfrak{g})$. Hence $\sigma(\bar{\theta}(z^{-1}))$ is the adjoint of $\sigma(z)$ $(z \varepsilon G_c)$. But $\bar{\theta}(N_c^-) = N_c^+$ and since ϕ_0 belongs to the highest weight, $\sigma(n')\phi_0 = \phi_0$ if $n' \varepsilon N_c^+$. Moreover from Lemma 13, $\Gamma(l_n w) = \Gamma(w)$ $(n \varepsilon N_c^-)$. Therefore if ψ_0 denotes the expression on the right hand side of the above

0149116

equation, $\psi_0(l_n w) = \psi_0(w)$ $(n \varepsilon N_c^-)$. Now let $a = \exp H \varepsilon \tilde{A}_c$ $(H \varepsilon \mathfrak{h})$. Then if $\bar{w} = \nu(w)$ $(w \varepsilon W)$,

$$(\phi_0, \sigma(\nu(l_a w))\phi_0) = (\sigma(\tilde{\theta}(\bar{a}^{-1}))\phi_0, \sigma(\bar{w})\phi_0) = e^{\Lambda_0(H)}(\phi_0, \sigma(\bar{w})\phi_0)$$

since Λ_0 is real. On the other hand it follows from Lemma 13 that $\lambda(\Gamma(l_a w)) = \lambda(H_+) + \lambda(\Gamma(w))$ where $H = H_+ + H'$ $(H_+ \varepsilon \mathfrak{c}_+, H' \varepsilon \mathfrak{h} \cap \mathfrak{m}')$. Therefore

$$\psi_0(l_a w) = \exp(\Lambda_0(H) + \lambda(H_+))\psi_0(w).$$

But $\Lambda(H) = \Lambda(H_+) + \Lambda(H')$ and it is obvious from Lemma 13 of [5(f)] that H' is a linear combination of H_{α_i} $(t < i \leqq l)$ so that $\lambda(H') = 0$. Hence

$$\Lambda_0(H) + \lambda(H_+) = \Lambda_0(H) + \lambda(H) = \Lambda(H).$$

This proves that $\psi_0(l_a w) = \xi(a)\psi_0(w)$ and so $\psi_0 \varepsilon \mathfrak{H}_\xi$. Moreover if $a \varepsilon A$,

$$\psi_0(r_a w) = (\phi_0, \sigma(\nu(w)\bar{a})\phi_0) \exp(\lambda(\Gamma(w)) + \lambda(\Gamma(a))) = \psi_0(w)\xi(a)$$

since $\sigma(\bar{a})\phi_0 e^{\lambda(\Gamma(a))} = \xi(a)\phi_0$. Hence

$$\psi_0(w^{\bar{h}}) = \psi_0(l_h r_{h^{-1}} w) = \xi(h)\xi(h^{-1})\psi_0(w) = \psi_0(w) \qquad (h \varepsilon A)$$

and therefore $\psi_0(w) = \int_{A_0} \psi_0(w^{\bar{h}})d\bar{h} = \psi_0(1)\psi(w)$ from Lemma 6. But $\psi_0(1) = |\phi_0|^2 = 1$ and so $\psi_0 = \psi$.

We have still to show that the space $\mathfrak{H}(\mu) \neq \{0\}$ for a suitable choice of μ. In order to do this we need some preliminary results on finite-dimensional representations of $\mathfrak{g}$ which we shall then apply to the above representation σ.

7. Some results on finite-dimensional representations. First we prove the following lemma.

LEMMA 15. *There exists an element $H \varepsilon \mathfrak{h}_0$ such that $\theta(X) = \exp(adH)X$ for all $X \varepsilon \mathfrak{g}$.*

Since $\theta(H) = H$ $(H \varepsilon \mathfrak{h})$ and $\theta^2(X) = X$ $(X \varepsilon \mathfrak{g})$, it is clear that $\theta(X_\alpha) = \pm X_\alpha$ for every root α. If $(\alpha_1, \cdots, \alpha_l)$ is a fundamental system of roots we can choose $H \varepsilon \mathfrak{h}$ such that $\theta(X_{\alpha_i}) = e^{\alpha_i(H)}X_{\alpha_i}$ $1 \leqq i \leqq l$. It is obvious that $\alpha_1(H), \cdots, \alpha_l(H)$ are all purely imaginary and therefore $H \varepsilon \mathfrak{h}_c$. Moreover since $[X_{\alpha_i}, X_{-\alpha_i}] \varepsilon \mathfrak{h}$, it follows that $\theta(X_{-\alpha_i}) = e^{-\alpha_i(H)}X_{\alpha_i}$. Therefore θ and $\exp(adH)$ are two automorphisms of $\mathfrak{g}$ which coincide on $\mathfrak{h}$ and also at X_{α_i}, $X_{-\alpha_i}$ $1 \leqq i \leqq l$. But $\mathfrak{g}$ is the smallest subalgebra of itself containing $\mathfrak{h}$ and X_{α_i}, $X_{-\alpha_i}$ $1 \leqq i \leqq l$ (see Lemmas 18 and 19 of [5(a)]) and so $\theta = \exp(adH)$.

LEMMA 16. *Let π be a representation of $\mathfrak{g}$ on a finite-dimensional space V. Then if ϕ is a vector belonging to some weight of π*

$$(\phi, \pi(z)\phi) = (\phi, \pi(\theta(z))\phi) \qquad (z \in G_c).$$

From Lemma 15 we can choose $h \in A_0$ such that $\theta(z) = hzh^{-1}$ for all $z \in G_c$. Then $(\phi, \pi(\theta(z))\phi) = (\pi(h^{-1})\phi, \pi(zh^{-1})\phi)$ since π is unitary on A_0. But since ϕ belongs to a weight of π, $\pi(h^{-1})\phi = c\phi$ where c is a unimodular complex number. Hence

$$(\phi, \pi(\theta(z))\phi) = |c|^2(\phi, \pi(z)\phi) = (\phi, \pi(z)\phi).$$

Now ϕ being as above, we propose to study the growth of the function $F(X) = (\phi, \pi(\exp X)\phi)$ $(X \in \mathfrak{p}_0)$ at infinity. Let X be a fixed element in $\mathfrak{p}_0$. Since $\theta(X) = -X$, $\pi(X)$ is a self-adjoint operator on V and therefore we can choose an orthonormal base $(\phi_1, \cdots, \phi_d)$ for V such that $\pi(X)\phi_i = \lambda_i\phi_i$ $(\lambda_i \in R,\ i = 1, \cdots, d)$. Then if $\phi = \sum_i a_i\phi_i$ $(a_i \in C)$,

$$F(X) = \sum_i |a_i|^2 e^{\lambda_i}.$$

Similarly $F(-X) = \sum_i |a_i|^2 e^{-\lambda_i}$. But since $\theta(X) = -X$, it follows from Lemma 16 that $F(X) = F(-X)$ and therefore $F(X) = \sum_i |a_i|^2 \cosh \lambda_i$. Now consider the function

$$c(t) = (\cosh t - 1)/t^2 = \sum_{n=1}^{\infty} t^{2n-2}/(2n)! \qquad (t \in R).$$

Then $c(t) = c(-t) \geqq 0$ and $c(t)$ increases with t for $t \geqq 0$. Also

$$F(X) = \sum_i |a_i|^2 + \sum_i |a_i\lambda_i|^2 c(\lambda_i) = 1 + \sum_i |a_i\lambda_i|^2 c(|\lambda_i|)$$

if we assume that $|\phi| = 1$. On the other hand $|\pi(X)\phi|^2 = \sum_i |a_i\lambda_i|^2$ and therefore $d^{-\frac{1}{2}} |\pi(X)\phi| \leqq \max_i |a_i\lambda_i|$. Let j be an index such that $|a_j\lambda_j| = \max_i |a_i\lambda_i|$. Then

$$\begin{aligned} F(X) &\geqq 1 + |a_j\lambda_j|^2 c(|\lambda_j|) \geqq 1 + |a_j\lambda_j|^2 c(|a_j\lambda_j|) \\ &\geqq 1 + d^{-1} |\pi(X)\phi|^2 c(d^{-\frac{1}{2}} |\pi(X)\phi|) = \cosh(d^{-\frac{1}{2}} |\pi(X)\phi|). \end{aligned}$$

Therefore we have the following result.

LEMMA 17. *Let π and ϕ be as in Lemma 16. Then if $|\phi| = 1$ and $X \in \mathfrak{p}_0$,*

$$(\phi, \pi(\exp X)\phi) \geqq \cosh(d^{-\frac{1}{2}} |\pi(X)\phi|)$$

where $d = \dim V$.

On the other hand we have the following result concerning $|\pi(X)\phi|$.

LEMMA 18. *Suppose ϕ belongs to the highest weight λ of π. Then if $\lambda(H_\beta) \neq 0$ for every noncompact root β, $|\pi(X)\phi|^2$ $(X \in \mathfrak{p}_0)$ is a positive definite quadratic form on $\mathfrak{p}_0$.*

Let X be an element in $\mathfrak{p}_0$. Then if we choose X_β, $X_{-\beta}$ as in [5(f), § 4], we can write $X = \sum_\beta (c_\beta X_\beta + \bar{c}_\beta X_{-\beta})$ where β runs over all noncompact positive roots and the bar denotes complex conjugate. Since ϕ belongs to the highest weight $\pi(X_\beta)\phi = 0$. Hence $\pi(X)\phi = \sum_\beta \bar{c}_\beta \pi(X_{-\beta})\phi$. On the other hand if $\lambda(H_\beta) \neq 0$ we know (Lemma 1 of [5(f)]) that $\pi(X_{-\beta})\phi \neq 0$ and it belongs to the weight $\lambda - \beta$. Since nonzero vectors belonging to distinct weights are linearly independent, we conclude that $\pi(X)\phi \neq 0$ unless $\bar{c}_\beta = 0$ for all β. This proves our assertion.

In Section 2 we have defined a (real) analytic mapping $z \to H(z)$ of G_c into $(-1)^{\frac{1}{2}}\mathfrak{h}_0$ such that $z \in U(\exp H(z))N_c^+$ $(z \in G_c)$. We shall say that a linear function λ on $\mathfrak{h}$ is *completely positive* if $\lambda(H_\alpha)$ is real and nonnegative for every positive root α.

LEMMA 19. *If λ is a completely positive linear function on $\mathfrak{h}$ then $\lambda(H(x)) \geqq 0$ for all $x \in G$. Moreover for any $X \in \mathfrak{p}_0$ the function $\lambda(H(\exp tX))$ $(t \in R)$ is non-decreasing for $t \geqq 0$. Finally $H(p) = H(p^{-1})$ if $p \in \exp \mathfrak{p}_0$.*

Let $\alpha_1, \cdots, \alpha_l$ be a fundamental system of positive roots. Define real linear functions $\Lambda_1, \cdots, \Lambda_l$ as follows: $\Lambda_i(H_{\alpha_j}) = \delta_{ij}$ $1 \leqq i, j \leqq l$ where $\delta_{ij} = 1$ or 0 according as $i = j$ or not. Then $\lambda = \lambda_1\Lambda_1 + \cdots + \lambda_l\Lambda_l$ where $\lambda_i = \Lambda(H_{\alpha_i})$ are nonnegative real numbers. Since Λ_i are dominant integral functions (see [5(a), p. 30]), it is clearly sufficient to prove the lemma under the assumption that λ is such a function. Let π be a finite-dimensional irreducible representation of $\mathfrak{g}$ on a space V with the highest weight λ [5(a), Theorem 1] and let ϕ be a unit vector in V belonging to the weight λ. Then if $x = uhn$ $(x \in G_0, u \in U, h \in A_+, n \in N_c^+)$ it is clear that

$$|\pi(x)\phi| = |\pi(uh)\phi| = |\pi(h)\phi| = e^{\lambda(H(x))}.$$

On the other hand $x = vp$ $(v \in K_0, p \in \exp \mathfrak{p}_0)$ and therefore

$$|\pi(x)\phi|^2 = |\pi(p)\phi|^2 = (\phi, \pi(p^2)\phi),$$

since $\pi(p)$ is self-adjoint. But it follows from Lemma 17 that $(\phi, \pi(p^2)\phi) \geqq 1$ and therefore $e^{\lambda(H(x))} = |\pi(x)\phi| \geqq 1$. This proves that $\lambda(H(x)) \geqq 0$. Moreover if $X \in \mathfrak{p}$, $\pi(X)$ is self-adjoint and therefore

$$\exp 2\lambda(H(\exp tX)) = |\pi(\exp tX)\phi|^2$$
$$= (\phi, \pi(\exp 2tX)\phi) = (\phi, \pi(\exp(-2tX)\phi) \qquad (t \in R)$$

from Lemma 16. Hence

$$\exp 2\lambda(H(\exp tX)) = \tfrac{1}{2}(\phi, \pi(\exp 2tX)\phi) + \tfrac{1}{2}(\phi, \pi(\exp(-2tX)\phi)$$
$$= \sum_{n \geqq 0} (2t)^{2n} |(\pi(X))^n \phi|^2 / 2n!.$$

Since $\lambda(H(\exp tX))$ is real and since only even powers of t occur and all the coefficients of the series are nonnegative, the second assertion of the lemma is now obvious. Also it is clear that $\lambda(H(p)) = \lambda(H(p^{-1}))$ for $p \in \exp \mathfrak{p}_0$. This is true in particular for $\lambda = \Lambda_i$ $i = 1, \cdots, l$. Since $\Lambda_1, \cdots, \Lambda_l$ is a base for the space of all linear functions on $\mathfrak{h}$, we conclude that $H(p) = H(p^{-1})$.

Let $B(X,Y) = sp(adXadY)$ $(X, Y \in \mathfrak{g})$. Then $-B(\tilde{\theta}(X), X)$ is a positive definite Hermitian form on $\mathfrak{g}$ which we denote by $\| X \|^2$. We now regard $\mathfrak{g}$ as a Helibert space under the corresponding norm $\| \cdot \|$. By combining Lemmas 17, 18 and 19 we can get the following stronger result.

LEMMA 20. *Suppose λ is completely positive and $\lambda(H_\beta) \neq 0$ for every noncompact root β. Then there exists a positive real number c such that $\lambda(H(\exp X)) \geqq c \|X\|$ for all $X \in \mathfrak{p}_0$ lying outside some bounded set.*

First suppose $\lambda(H_{\alpha_i})$ $1 \leqq i \leqq l$ are all integers so that λ is a dominant integral function. We define π, V and ϕ as in the proof of Lemma 19. Then from Lemma 17,

$$(\phi, \pi(\exp 2X)\phi) \geqq \cosh(2d^{-\frac{1}{2}} |\pi(X)\phi|)$$

where $d = \dim V$. But we have seen above that if $p = \exp X$,

$$(\phi, \pi(\exp 2X)\phi) = |\pi(p)\phi|^2 = e^{2\lambda(H(p))}.$$

Hence

$$\lambda(H(p)) \geqq \tfrac{1}{2}\log(\cosh(2d^{-\frac{1}{2}}|\pi(X)\phi|)) \geqq d^{-\frac{1}{2}}|\pi(X)\phi| - \tfrac{1}{2}\log 2.$$

On the other hand we know from Lemma 18 that $|\pi(X)\phi|^2$ $(X \in \mathfrak{p}_0)$ is a positive definite quadratic form on $\mathfrak{p}_0$. Hence if c_0 is the least possible value of $|\pi(X)\phi|$ for all $X \in \mathfrak{p}_0$ with $\| X \| = 1$, $c_0 > 0$ and $|\pi(X)\phi| \geqq c_0 \| X \|$ $(X \in \mathfrak{p}_0)$. Therefore

$$\lambda(H(\exp X)) \geqq c_0 d^{-\frac{1}{2}} \| X \| - \tfrac{1}{2}\log 2 \geqq \tfrac{1}{2} c_0 d^{-\frac{1}{2}} \| X \|$$

provided $\| X \| \geqq c_0^{-1} d^{\frac{1}{2}} \log 2$ $(X \in \mathfrak{p}_0)$. Hence we may take $c = \tfrac{1}{2} c_0 d^{-\frac{1}{2}}$.

Now we come to the general case. Define $\Lambda_1, \cdots, \Lambda_l$ as in the proof of Lemma 19. Then $\lambda = \lambda_1\Lambda_1 + \cdots + \lambda_l\Lambda_l$ where $\lambda_i = \lambda(H_{\alpha_i})$ are non-negative real numbers. Choose a positive number r so large that if $\lambda_i \neq 0$, $r\lambda_i \geqq 1$ $(1 \leqq i \leqq l)$ and put $\lambda_0 = a_1\Lambda_1 + \cdots + a_l\Lambda_l$ where $a_i = 0$ or 1 according as $\lambda_i = 0$ or not. Then $r\lambda - \lambda_0$ is completely positive and λ_0 is a dominant integral function. Let α be a positive root. Then $H_\alpha = b_1H_{\alpha_1} + \cdots + b_lH_{\alpha_l}$ $(b_i \varepsilon R, b_i \geqq 0)$. Suppose $b_i > 0$ $(1 \leqq i \leqq m)$ and $b_i = 0$ $(m < i \leqq l)$. Then $\lambda(H_\alpha) = b_1\lambda_1 + \cdots + b_m\lambda_m$ and $\lambda_0(H_\alpha) = b_1a_1 + \cdots + b_ma_m$. Since λ_i, a_i are all nonnegative and since $\lambda_i = 0$ if and only if $a_i = 0$, it is clear that $\lambda(H_\alpha) = 0$ if and only if $\lambda_0(H_\alpha) = 0$. Therefore $\lambda_0(H_\beta) \neq 0$ for any noncompact positive root β and so by the above proof, there exists a positive number c' such that $\lambda_0(H(\exp X)) \geqq c' \| X \|$ $(X \varepsilon \mathfrak{p}_0)$ provided that $\| X \|$ is sufficiently large. However since $r\lambda - \lambda_0$ is completely positive

$$r\lambda(H(\exp X)) \geqq \lambda_0(H(\exp X))$$

(Lemma 19) and therefore

$$\lambda(H(\exp X)) \geqq c' \| X \| / r \qquad (X \varepsilon \mathfrak{p}_0)$$

if $\| X \|$ is sufficiently large.

On the other hand it is quite easy to obtain a result in the opposite direction.

LEMMA 21. *Let λ be a linear function on $\mathfrak{h}$. Then there exists a real number c' such that $|\lambda(H(\exp X))| \leqq c' \| X \|$ for all $X \varepsilon \mathfrak{p}_0$.*

Put $\lambda_i = \Lambda(H_{\alpha_i})$ $i = 1, 2, \cdots, l$. Then $\lambda = \lambda_1\Lambda_1 + \cdots + \lambda_l\Lambda_l$ and $|\lambda(H)| \leqq \sum_i |\lambda_i| \, |\Lambda_i(H)|$ $(H \varepsilon \mathfrak{h})$. Therefore it is obviously sufficient to prove the lemma under the assumption that λ is a dominant integral function. Define π, V and ϕ as above. Then if $p = \exp X$ $(X \varepsilon \mathfrak{p}_0)$,

$$e^{\lambda(H(p))} = |\pi(p)\phi| \leqq |\pi(p)| \leqq e^{|\pi(X)|}$$

where $|T|$ denotes, as usual, the bound of an operator T. Therefore from Lemma 19,

$$|\lambda(H(p))| = \lambda(H(p)) \leqq |\pi(X)| \leqq \{sp(\pi(X))^2\}^{\frac{1}{2}}$$

since $\pi(X)$ is self-adjoint. If a' is the maximum value of $sp(\pi(X))^2$ for all $X \varepsilon \mathfrak{p}_0$ with $\| X \| = 1$, it is clear that $sp(\pi(X))^2 \leqq a' \| X \|^2$ for every X in $\mathfrak{p}_0$ and therefore $|\lambda(H(\exp X))| \leqq c' \| X \|$ where $c' = (a')^{\frac{1}{2}}$.

We shall also need another similar result. Let ρ be defined as in Lemma 3.

LEMMA 22. *Let π be a representation of $\mathfrak{g}$ on a finite-dimensional space V. Then there exist positive real numbers a, b such that*

$$|\pi(x^{-1})\phi| \leqq be^{a\rho(H(x))}|\phi|$$

for every $\phi \varepsilon V$ and $x \varepsilon G_0$.

If $x = up$ $(u \varepsilon K_0, p \varepsilon \exp \mathfrak{p}_0)$, $\pi(x^{-1})\phi = \pi(p^{-1})\phi'$ where $\phi' = \pi(u)\phi$. Also $H(x) = H(p)$ and $|\phi'| = |\phi|$. So it is enough to consider the case when $u = 1$. Let $p = \exp X$ $(X \varepsilon \mathfrak{p}_0)$. Then

$$|\pi(p^{-1})\phi| \leqq |\pi(p^{-1})|\,|\phi| \leqq e^{|\pi(X)|}|\phi|.$$

But as we have just seen above, there exists a real number a_1 such that $|\pi(X)| \leqq a_1 \| X \|$ for all $X \varepsilon \mathfrak{p}_0$. On the other hand from Lemma 20 we can find positive real numbers a_2 and M such that[10]

$$\rho(H(\exp X)) \geqq a_2 \| X \|$$

for all $X \varepsilon \mathfrak{p}_0$ for which $\| X \| \geqq M$. Let b be an upper bound for $|\pi(\exp(-X))|$ when $\| X \| \leqq M$ $(X \varepsilon \mathfrak{p}_0)$. Then if $a = a_1/a_2$ it is obvious that $|\pi(x^{-1})\phi| \leqq be^{a\rho(H(x))}$ for all $\phi \varepsilon V$ and $x \varepsilon G$.

The following lemma describes some properties of the mapping $x \to \Gamma(x)$ $(x \varepsilon G)$.

LEMMA 23. *Let λ be a real linear function on $\mathfrak{h}$. Then if $u \varepsilon K$, the real part of $\lambda(\Gamma(u))$ is zero. Moreover if $\lambda(H_\alpha) = 0$ for every positive root α which is not totally positive,*

$$\lambda(\Gamma(\exp X)) = 2\lambda(H(\exp \tfrac{1}{2}X)) \qquad (X \varepsilon \mathfrak{p}_0).$$

We know that $\mathfrak{h}$ is the direct sum of $\mathfrak{c}_+$ and $\mathfrak{h} \cap \mathfrak{m}'$ (in the notation of Section 6) and if α is a positive root which is not totally positive, $H_\alpha \varepsilon \mathfrak{h} \cap \mathfrak{m}'$. Without affecting its values on $\mathfrak{c}_+$ we can replace λ by a linear function which vanishes identically on $\mathfrak{h} \cap \mathfrak{m}'$ and therefore assume that $\lambda(H_\alpha) = 0$ for all positive roots α which are not totally positive. Then $\lambda = \lambda_1\Lambda_1 + \cdots + \lambda_t\Lambda_t$ (in the notation of the proof of Lemma 14) where $\lambda_1, \cdots, \lambda_t$ are real numbers. Obviously it would be enough to prove the lemma for $\lambda = \Lambda_i$ $i = 1, \cdots, t$. Hence we may assume that λ is a dominant integral function. Let π be an irreducible representation of $\mathfrak{g}$ on a finite-dimensional space V with the highest weight λ and let ϕ be a unit vector in V belonging to the weight λ. Then if β is a positive root which is not totally positive, $\lambda(H_\beta) = 0$ and therefore from Lemma 1 of [5(f)] $\pi(X_{-\beta})\phi = 0$. Hence it is clear that

[10] Here we have to make use of the fact (see Weyl [9]) that $\rho(H_\alpha) > 0$ for every positive root α.

$\pi(\mathfrak{m}'+\mathfrak{p}_+)\phi=\{0\}$. On the other hand we have seen in Section 6 that $\bar{x}=qm(\bar{x})p$ $(x\,\varepsilon\,G)$ where $q\,\varepsilon\,P_c^-$, $p\,\varepsilon\,P_c^+$. Therefore

$$(\phi,\pi(\bar{x})\phi)=(\pi(\dot{\theta}(q^{-1}))\phi,\pi(m(\bar{x}))\phi)=(\phi,\pi(m(\bar{x})\phi)$$

since $\bar{\theta}(q^{-1})\,\varepsilon\,P_c^+$. Moreover since $\pi(\mathfrak{m}')\phi=\{0\}$, it is obvious that

$$\pi(m(\bar{x}))\phi=e^{\lambda(\Gamma(x))}\phi$$

and therefore $(\phi,\pi(\bar{x})\phi)=e^{\lambda(\Gamma(x))}$ $(x\,\varepsilon\,G)$. Now if $u\,\varepsilon\,K$, $\bar{u}\,\varepsilon\,M_c$ and therefore $\pi(\bar{u})\phi=e^{\lambda(\Gamma(u))}\phi$. But since $\pi(\bar{u})$ is unitary, $\lambda(\Gamma(u))$ must be purely imaginary. On the other hand if $x=p^2$ where $p=\exp\frac{1}{2}X$ $(X\,\varepsilon\,\mathfrak{p}_0)$, the above equation gives $|\pi(\bar{p})\phi|^2=e^{\lambda(\Gamma(p^2))}$. But we have seen that $|\pi(\bar{p})\phi|=e^{\lambda(H(\bar{p}))}$. Hence

$$\lambda(\Gamma(\exp X))-2\lambda(H(\exp\tfrac{1}{2}X))$$

must be an integral multiple of $2\pi(-1)^{\frac{1}{2}}$. However it is a continuous function of X and it is zero at $X=0$. Therefore since $\mathfrak{p}_0$ is connected it must be everywhere zero on $\mathfrak{p}_0$. This proves the lemma.

COROLLARY. *If $u\,\varepsilon\,K$, $\Gamma(u)$ lies in $\mathfrak{h}_0$. Moreover*

$$\Gamma(\exp X)-\tfrac{1}{2}H(\exp\tfrac{1}{2}X)\,\varepsilon\,\mathfrak{h}\cap\mathfrak{m}'$$

for all $X\,\varepsilon\,\mathfrak{p}_0$.

This is an immediate consequence of the above lemma.

8. Proof of the existence of representations. In order to show that the space $\mathfrak{H}_\Lambda(\mu)$ of Lemma 9 is not zero for some μ, it is enough to find a function μ on G_0, satisfying the conditions of Section 6 and such that

$$\int_{G_0}|\psi(x)|^2\mu(\bar{x})\,d\bar{x}<\infty$$

where ψ is the function of Lemma 14. But $|\psi(x^{-1})|\leqq|\sigma(\bar{x}^{-1})\phi_0|\,e^{\Re(\lambda(\Gamma(x^{-1})))}$, (in the notation of Lemma 14) where $\Re c$ denotes the real part of a complex number c. On the other hand we know from Lemma 29 that

$$|\sigma(\bar{x}^{-1})\phi_0|\leqq b e^{a\rho(H(\bar{x}))}\qquad(x\,\varepsilon\,G)$$

for suitable real numbers a and b. If $x=up$ $(u\,\varepsilon\,K,\ p=\exp X,\ X\,\varepsilon\,\mathfrak{p}_0)$ $\Gamma(x^{-1})=\Gamma(p^{-1})+\Gamma(u^{-1})$ (Lemma 13) and therefore

$$\Re(\lambda(\Gamma(x^{-1})))=\lambda(\Gamma(p^{-1}))=2\lambda(H(\exp(-\tfrac{1}{2}X)))=2\lambda(H(\exp\tfrac{1}{2}X))$$

from Lemmas 23 and 19. It is obvious[10] that $r'\rho-2\lambda$ is completely positive for a suitable real $r'\geqq 0$. Then

$$2\lambda(H(\exp\tfrac{1}{2}X))\leqq r'\rho(H(\exp\tfrac{1}{2}X))\leqq r'\rho(H(\exp X))=r'\rho(H(\bar{x}))$$

from Lemma 19. Hence $|\psi(x^{-1})| \leq b e^{(a+r')\rho(H(\bar{x}))}$ $(x \in G)$. Now put $\mu(\bar{x}) = e^{-r\rho(H(\bar{x}^{-1}))}$ $(\bar{x} \in G_0)$ where $r \geq 2a + 2r' + 4$. Then μ is a continuous function on G_0 which is everywhere positive and

$$\int_{G_0} |\psi(x)|^2 \mu(\bar{x}) d\bar{x} = \int_{G_0} |\psi(x^{-1})|^2 \mu(\bar{x}^{-1}) d\bar{x} \leq b^2 \int_{G_0} e^{-4\rho(H(\bar{x}))} d\bar{x} < \infty$$

from Lemma 3. Moreover since ρ is a dominant integral function (see Weyl [9]), there exists a finite-dimensional irreducible representation τ of $\mathfrak{g}$ with the highest weight ρ. Let ζ be a unit vector in the representation space belonging to the weight ρ. Then $|\tau(\bar{x}^{-1})\zeta| = e^{\rho(H(\bar{x}^{-1}))}$ $(\bar{x} \in G_0)$ and therefore, if $\bar{x}, \bar{y} \in G_0$,

$$e^{\rho(H(\bar{y}^{-1}\bar{x}^{-1}))} = |\tau(\bar{y}^{-1}\bar{x}^{-1})\zeta| \leq |\tau(\bar{y}^{-1})| \, e^{\rho(H(\bar{x}^{-1}))}.$$

This proves that $\mu(\bar{x}\bar{y}) \leq |\tau(\bar{y})|^r \mu(\bar{x})$ and since $|\tau(\bar{y})|$ is bounded on every compact set in G_0, μ fulfills all the conditions of Section 6. Hence $\psi \in \mathfrak{H}_\Lambda(\mu)$ and so $\mathfrak{H}_\Lambda(\mu) \neq \{0\}$.

We can now summarise our results in the following theorem.

THEOREM 2. *Let Λ be a real linear function on $\mathfrak{h}$ such that $\Lambda(H_\alpha)$ is a nonnegative integer for every positive root α which is not totally positive. Let ξ denote the character of $\tilde{A}_c$ defined by $\xi(\exp H) = e^{\Lambda(H)}$ $(H \in \mathfrak{h})$ and let μ be a measurable function on G_0 which is everywhere positive and such that $\sup_{\bar{x} \in G_0} \mu(\bar{x}\bar{y})/\mu(\bar{x})$ $(\bar{y} \in G_0)$ remains bounded on every compact subset of G_0. Let $\mathfrak{H}_\Lambda(\mu)$ denote the Hilbert space of all holomorphic functions f on W which fulfill the following two conditions:*

(1) $f(l_{na}w) = \xi(a)f(w)$ $\quad (a \in \tilde{A}_c, n \in N_c^-, w \in W)$

(2) $\|f\|^2 = \int_{G_0} |f(x)|^2 \mu(\bar{x}) d\bar{x} < \infty.$

Then if $\mathfrak{H}_\Lambda(\mu) \neq \{0\}$ there exists a unique function ψ in $\mathfrak{H}_\Lambda(\mu)$ such that $\psi(1) = 1$ and $\psi(w^{\bar{h}}) = \psi(w)$ for all $\bar{h} \in A_0$ and $w \in W$. Put $\psi_x(w) = \psi(r_x w)$ $(x \in G, w \in W)$ and let $\mathfrak{H}$ denote the smallest closed subspace of $\mathfrak{H}_\Lambda(\mu)$ containing ψ_x for all $x \in G$. Then we can define a quasi-simple irreducible representation π of G on $\mathfrak{H}$ by the rule

$$\pi(x)\phi(w) = \phi(r_x w) \qquad (x \in G, \phi \in \mathfrak{H}, w \in W).$$

This representation is infinitesimally equivalent to the representation π_Λ of [5(f), § 6]. *Finally, it is always possible to choose μ in such a way that $\mathfrak{H}_\Lambda(\mu) \neq \{0\}$.*

The infinitesimal equivalence follows immediately from Lemmas 8 and 10, Theorem 2 of [5(f)] and Theorem 5 of [5(b)].

COROLLARY. *Suppose* $\mu = 1$ *and* $\mathfrak{H}_\Lambda(\mu) \neq \{0\}$. *Then* $\mathfrak{H} = \mathfrak{H}_\Lambda(\mu)$, π *is unitary and*

$$(\psi, \pi(x)\psi) = \psi(x) \| \psi \|^2 \qquad (x \varepsilon G).$$

It is obvious that if $f \varepsilon \mathfrak{H}_\Lambda(\mu)$,

$$\int_{G_0} | f(xy) |^2 d\bar{x} = \| f \|^2 \qquad (y \varepsilon G),$$

and therefore from Lemma 12, $\mathfrak{H} = \mathfrak{H}_\Lambda(\mu)$ and π is unitary. Moreover

$$\int_{A_0} \psi(y^{\bar{h}} x) d\bar{h} = \psi(x)\psi(y) \qquad (x, y \varepsilon G)$$

from Lemma 6. Hence[11]

$$\psi(x) \| \psi \|^2 = \int_{G_0} \{ (\operatorname{conj} \psi(y)) \int_{A_0} \psi(y^{\bar{h}} x) d\bar{h} \} d\bar{y}$$
$$= \int_{A_0} d\bar{h} \int_{G_0} (\operatorname{conj} \psi(y)) \psi(yx) d\bar{y} = (\psi, \pi(x)\psi)$$

by Fubini's Theorem.

9. Unitary representations. Our next object is to look for unitary representations among those constructed above. However the result which we give below is not quite as strong as Theorem 3 of [5(f)] which was stated there without proof. We have to postpone its proof to another paper since it requires a deeper study of the representations.

THEOREM 3. *Let* $(\alpha_1, \cdots, \alpha_l)$ *be a fundamental system of positive roots and suppose* $(\alpha_1, \cdots, \alpha_t)$ *are all the totally positive roots in this system. Let* λ_i $(t < i \leqq l)$ *be given nonnegative integers such that* $\lambda_i = 0$ *if* α_i *is not a root of*[3] $\mathfrak{g}_+$. *Then we can find a real number* c *with the following property. If* Λ *is a real linear function on* $\mathfrak{h}$ *and* $\Lambda(H_{\alpha_i}) \leqq c$ $(1 \leqq i \leqq t)$, $\Lambda(H_{\alpha_i}) = \lambda_i$ $(t < i \leqq l)$, *then*[3] π_Λ *is infinitesimally unitary.*

Let Λ be a real linear function such that $\Lambda(H_{\alpha_i}) = \lambda_i$ $(t < i \leqq l)$. Then $\Lambda(H_\alpha) = 0$ for every root α of $\mathfrak{g}'$ (see [5(f), § 5]) and therefore $\pi_\Lambda(X) = 0$ if $X \varepsilon \mathfrak{g}'$ (Lemma 19 of [5(f)]). Since $\mathfrak{g}_0 = \mathfrak{g}_+ \cap \mathfrak{g}_0 + \mathfrak{g}' \cap \mathfrak{g}_0$ we can now restrict our attention to $\mathfrak{g}_+ \cap \mathfrak{g}_0$. Hence without any essential loss of generality, we may asume that $\mathfrak{g}' = \{0\}$ and so $\mathfrak{g} = \mathfrak{g}_+$. Then it follows

---

[11] conj c denotes the complex conjugate of a number $c \varepsilon C$.

from Lemma 13 of [5(f)] that α_i $(t < i \leqq l)$ are all compact. Define a linear function λ on $\mathfrak{h}$ by the conditions $\lambda(H_{\alpha_i}) = \Lambda(H_{\alpha_i})$ $1 \leqq i \leqq t$ and $\lambda(H_{\alpha_i}) = 0$ $t < i \leqq l$. Then λ is real and $\Lambda_0 = \Lambda - \lambda$ is a dominant integral function on $\mathfrak{h}$ and

$$\psi(x) = (\phi_0, \sigma(\bar{x})\phi_0)e^{\lambda(\Gamma(x))} \qquad (x \in G)$$

in the notation of Lemma 14. Now consider the integral

$$\int_{G_0} |\psi(x)|^2 \, d\bar{x}.$$

We know that

$$|\psi(x)| \leqq |\sigma(\bar{x})\phi_0| \, e^{\Re(\lambda(\Gamma(x)))} = \exp(\Lambda_0(H(\bar{x})) + \Re(\lambda(\Gamma(x))).$$

But if $x = u \exp X$ $(u \in K, X \in \mathfrak{p}_0)$

$$\Re(\lambda(\Gamma(x))) = \lambda(\Gamma(\exp X)) = 2\lambda(H(\exp \tfrac{1}{2}X))$$

from Lemmas 13 and 23. Now define Λ_i $1 \leqq i \leqq l$ as in the proof of Lemma 19. Then $\lambda_0 = \Lambda_1 + \cdots + \Lambda_t$ is a completely positive linear function and it follows from Lemma 13 of [5(f)] that $\lambda_0(H_\beta) = 1$ for every totally positive root. Since $\mathfrak{g} = \mathfrak{g}_+$, every noncompact positive root is totally positive and therefore Lemma 20 is applicable to λ_0. Hence there exists a positive number a such that

$$\lambda_0(H(\exp X)) \geqq a \, \| X \|,$$

for all $X \in \mathfrak{p}_0$ lying outside some bounded set. On the other hand from Lemma 21 we can find a real constant $b \geqq 0$ such that

$$\Lambda_0(H(\exp X)) + 2\rho(H(\exp X)) \leqq b \, \| X \| \qquad (X \in \mathfrak{p}_0).$$

Hence $\Lambda_0(H(\exp X)) + 2\rho(H(\exp X)) \leqq (2b/a)\lambda_0(H(\exp \tfrac{1}{2}X))$, for all $X \in \mathfrak{p}_0$ with a sufficiently large value of $\| X \|$. We note that b depends only on Λ_0 and therefore only on the integers λ_i $(t < i \leqq l)$. Now put $c = -b/a$. Then if $\Lambda(H_{\alpha_i}) \leqq c$ $1 \leqq i \leqq t$, $-\lambda + c\lambda_0$ is completely positive and therefore from Lemma 19

$$-2\lambda(H(\exp \tfrac{1}{2}X)) \geqq -2c\lambda_0(H(\exp \tfrac{1}{2}X)) \geqq \Lambda_0(H(\exp X)) + 2\rho(H(\exp X)),$$

if $\| X \|$ is sufficiently large $(X \in \mathfrak{p}_0)$. Hence

$$\exp\{\Lambda_0(H(\exp X)) + 2\rho(H(\exp X)) + \lambda(\Gamma(\exp X))\}$$

is bounded on $\mathfrak{p}_0$ and if M is an upper bound for it, it is obvious that

$$|\psi(x)| \leqq M e^{-2\rho(H(\bar{x}))} \qquad (x \in G).$$

Therefore $\int_{G_0} |\psi(x)|^2 d\bar{x} \leqq M^2 \int_{G_0} e^{-4\rho(H(\bar{x}))} d\bar{x} < \infty$ from Lemma 3. This proves that the space $\mathfrak{H}_\Lambda = \mathfrak{H}_\Lambda(1)$ (corresponding to the function $\mu = 1$) is not zero if $\Lambda(H_{\alpha_i}) \leqq c$ $(1 \leqq i \leqq t)$ and our assertion now follows immediately from Theorem 2 and its corollary.

10. A result on characters. Let Λ_0 be a linear function on $\mathfrak{h}$ satisfying the conditions of Theorem 2. We denote by T_{Λ_0} the character [5(c)] of the quasi-simple irreducible representation of G defined in Theorem 2 corresponding to Λ_0. Since two infinitesimally equivalent quasi-simple irreducible representations have the same character (see [5(c), § 7]), T_{Λ_0} is independent of the choice of μ in Theorem 2 and so it is completely determined by Λ_0. We shall now try to obtain some information about this character under suitable assumptions on Λ_0.

Let $\mathfrak{H}$ be a Hilbert space and Q a bounded operator on $\mathfrak{H}$. We say that Q is *summable* if there exists a complete orthonormal set $(\psi_j)_{j \varepsilon J}$ in $\mathfrak{H}$ and a regular operator B such that $\sum_{i,j \varepsilon J} |q_{ij}| < \infty$ where $q_{ij} = (\psi_i, BQB^{-1}\psi_j)$. Let $C_c^\infty(A)$ denote the set of all complex-valued functions on A which are everywhere indefinitely differentiable and which vanish outside a compact set. Since K is simply connected,[4] there exists (see Weyl [9]) an analytic function Δ_k on A such that

$$\Delta_k(\exp H) = \prod_{\alpha \varepsilon P_k} (e^{\alpha(H)/2} - e^{-\alpha(H)/2}) \qquad (H \varepsilon \mathfrak{h}_0)$$

where P_k is the set of all positive compact roots of $\mathfrak{g}$. As before we write $y^{\bar{x}} = xyx^{-1}$ $(x, y \varepsilon G)$. Let dh and $d\bar{u}$ denote the Haar measures on A and K_0 respectively. (We assume that $\int_{K_0} d\bar{u} = 1$). Then we have the following lemma.

LEMMA 24. *Let π be a quasi-simple irreducible representation of G on $\mathfrak{H}$. Then if $f \varepsilon C_c^\infty(A)$, the operator*

$$\int_A \int_{K_0} f(h) \Delta_k(h) \pi(h^{\bar{u}}) \, dh d\bar{u}$$

is summable.

We can choose a homomorphism η of K into C such that $\eta(u^{-1})\pi(u)$ $(u \varepsilon K)$ depends only on $\bar{u}$ (see [5(c), p. 249]). Hence if we put $\bar{\pi}(\bar{u}) = \eta(u^{-1})\pi(u)$ it is clear that $\bar{\pi}$ is a representation of the compact group K_0. Therefore we can find a regular operator B such that $B\bar{\pi}(\bar{u})B^{-1}$ is unitary for every $\bar{u} \varepsilon K_0$. Hence, in view of our definition of the summa-

bility of an operator, it is clear that, without any loss of generality, we may assume that $\bar{\pi}$ itself is a unitary representation. Ω being the set of all equivalence classes of irreducible finite-dimensional representations of K, we denote, as usual, by $\mathfrak{H}_{\mathfrak{D}}$ ($\mathfrak{D} \in \Omega$) the subspace of $\mathfrak{H}$ consisting of those elements which transform under $\pi(K)$ according to $\mathfrak{D}$. Then there exists an integer N such that $\dim \mathfrak{H}_{\mathfrak{D}} \leqq N(d(\mathfrak{D}))^2$ ($\mathfrak{D} \in \Omega$) where $d(\mathfrak{D})$ is the degree of any representation in $\mathfrak{D}$ (see [5(c), Theorem 4]). Since $\bar{\pi}$ is unitary, the subspaces $\mathfrak{H}_{\mathfrak{D}}$ are mutually orthogonal and so we can choose a complete orthonormal set $(\psi_j)_{j \in J}$ in $\mathfrak{H}$ such that each ψ_j lies in some $\mathfrak{H}_{\mathfrak{D}}$ (see [5(a), Theorem 4]). Let $J(\mathfrak{D})$ denote the set of all $j \in J$ for which $\psi_j \in \mathfrak{H}_{\mathfrak{D}}$. Then $(\psi_j)_{j \in J(\mathfrak{D})}$ is an orthonormal base for $\mathfrak{H}_{\mathfrak{D}}$. Put $n(\mathfrak{D}) = (d(\mathfrak{D}))^{-1} \dim \mathfrak{H}_{\mathfrak{D}}$. Then $n(\mathfrak{D})$ is a nonnegative integer. Also if $E_{\mathfrak{D}}$ is the orthogonal projection of $\mathfrak{H}$ on $\mathfrak{H}_{\mathfrak{D}}$, it follows from the Schur orthogonality relations on the compact group K_0 that

$$E_{\mathfrak{D}}\Big(\int_{K_0} \pi(h^{\bar{u}})\,d\bar{u}\Big)E_{\mathfrak{D}} = \eta(h)E_{\mathfrak{D}}\Big(\int_{K_0} \bar{\pi}(\bar{u}\bar{h}\bar{u}^{-1})\,d\bar{u}\Big)E_{\mathfrak{D}}$$
$$= d(\mathfrak{D})^{-1}\, sp(E_{\mathfrak{D}}\pi(h)E_{\mathfrak{D}})E_{\mathfrak{D}} = d(\mathfrak{D})^{-1}\zeta_{\mathfrak{D}}(h)E_{\mathfrak{D}} \qquad (h \in A)$$

where $\zeta_{\mathfrak{D}}$ is the character (on K) of the class $\mathfrak{D}$. This shows that if $i, j \in J(\mathfrak{D})$,

$$(\psi_i, Q_f\psi_j) = d(\mathfrak{D})^{-1}\delta_{ij}\int_A f(h)\Delta_k(h)\zeta_{\mathfrak{D}}(h)\,dh$$

where δ_{ij} is the Kronecker symbol and Q_f is the operator of our lemma. Therefore if Ω_π is the set of all $\mathfrak{D} \in \Omega$ such that $\mathfrak{H}_{\mathfrak{D}} \neq \{0\}$,

$$\sum_{i,j \in J} |(\psi_i, Q_f\psi_j)| \leqq N \sum_{\mathfrak{D} \in \Omega_\pi} d(\mathfrak{D}) \Big| \int_A f(h)\Delta_k(h)\zeta_{\mathfrak{D}}(h)\,dh \Big|.$$

Let $\mathfrak{D}$ be a class in Ω_π and σ a representation in $\mathfrak{D}$. We denote the corresponding representation of $\mathfrak{k}$ also by σ. Since $\mathfrak{k} = \mathfrak{k}' + \mathfrak{c}$, it follows from Schur's lemma that the representation space V of σ is irreducible under $\sigma(\mathfrak{k}')$. Let σ' denote the corresponding representation of $\mathfrak{k}'$ on V. Since $\mathfrak{h}$ is the direct sum of $\mathfrak{c}$ and $\mathfrak{h}' = \mathfrak{h} \cap \mathfrak{k}'$ we may identify linear functions on $\mathfrak{h}'$ with those linear functions on $\mathfrak{h}$ which vanish identically on $\mathfrak{c}$. Now $\mathfrak{k}'$ is semisimple and $\mathfrak{h}'$ is a Cartan subalgebra of $\mathfrak{k}'$ and it is clear that the roots of $\mathfrak{k}'$ with respect to $\mathfrak{h}'$ then coincide with the compact roots of $\mathfrak{g}$. Let $\lambda_{\mathfrak{D}}$ denote the highest weight of σ' (if we take P_k as the set of positive roots of $\mathfrak{k}'$). Then if $2\rho_k = \sum_{\alpha \in P_k} \alpha$ we know (see Weyl [9]) that

$$\Delta_k(\exp H)\zeta_{\mathfrak{D}}(\exp H) = \eta(\exp H)\sum_{s \in \mathfrak{w}_k} \epsilon(s)e^{s\lambda'_{\mathfrak{D}}(H)} \qquad (H \in \mathfrak{h}' \cap \mathfrak{k}_0)$$

where $\lambda'_{\mathfrak{D}} = \lambda_{\mathfrak{D}} + \rho_k$ and $\mathfrak{w}_k$ is the group generated by the Weyl reflexions s_α corresponding to $\alpha \varepsilon P_k$, and $\epsilon(s) = (-1)^r$ where r is the number of negative roots among $s\alpha$ $(\alpha \varepsilon P_k)$. On the other hand the degrees of σ and σ' are the same and therefore (Weyl [9])

$$d(\mathfrak{D}) = \prod_{\alpha \varepsilon P_k} \lambda'_{\mathfrak{D}}(H_\alpha)/\rho_k(H_\alpha).$$

Now if we regard η as a representation of K of degree 1, it is obvious that $\eta(\exp X) = e^{\mu(X)}$ $(X \varepsilon \mathfrak{k}_0)$ where μ is a linear function on $\mathfrak{k}$ which vanishes identically on $\mathfrak{k}'$. Select a base $\Gamma_1, \cdots, \Gamma_r$ for $\mathfrak{c}_0 = \mathfrak{c} \cap \mathfrak{k}_0$ (over R) such that $\exp(t_1\Gamma_1 + \cdots + t_r\Gamma_r) = 1$ in K_0 $(t_1, \cdots, t_r \varepsilon R)$ if and only if $t_1, \cdots, t_r$ are all integers (see [5(c), p. 239]). Then since $\eta(h^{-1})\sigma(h)$ $(h \varepsilon A)$ depends only on $\bar{h}$, it follows that $\sigma(\Gamma_i) = \mu(\Gamma_i) + 2\pi(-1)^{\frac{1}{2}}m_i$ $1 \leqq i \leqq r$ where m_i are integers. We denote by $m_{\mathfrak{D}}$ the linear function on $\mathfrak{h}$ given by $m_{\mathfrak{D}}(H) = 0$ $(H \varepsilon \mathfrak{h}')$ and $m_{\mathfrak{D}}(\Gamma_i) = 2\pi(-1)^{\frac{1}{2}}m_i$ $1 \leqq i \leqq r$. Also put

$$| m_{\mathfrak{D}} | = (m_1^2 + \cdots + m_r^2 + 1)^{\frac{1}{2}}.$$

Then if $\nu_{\mathfrak{D}} = \lambda_{\mathfrak{D}} + m_{\mathfrak{D}} + \rho_k$,

$$\Delta_k(\exp H)\zeta_{\mathfrak{D}}(\exp H) = \eta(\exp H) \sum_{s \varepsilon \mathfrak{w}_k} \epsilon(s) e^{s\nu_{\mathfrak{D}}(H)} \qquad (H \varepsilon \mathfrak{h}_0)$$

and

$$d(\mathfrak{D}) = \prod_{\alpha \varepsilon P_k} \nu_{\mathfrak{D}}(H_\alpha)/\rho_k(H_\alpha)$$

since $sm_{\mathfrak{D}} = m_{\mathfrak{D}}$ $(s \varepsilon \mathfrak{w}_k)$ and $H_\alpha \varepsilon \mathfrak{h}'$ $(\alpha \varepsilon P_k)$.

Let $\mathfrak{U}$ be the subalgebra of $\mathfrak{B}$ generated by $(1, \mathfrak{h})$. We regard elements of $\mathfrak{U}$ as differential operators on A in the usual way so that

$$(Hg)(h) = \{(d/dt)g(h \exp tH)\}_{t=0} \qquad (h \varepsilon A, H \varepsilon \mathfrak{h}_0, t \varepsilon R, g \varepsilon C_c^\infty(A)).$$

Since any element $s \varepsilon \mathfrak{w}_k$ permutes the compact roots among themselves, it follows that

$$\prod_{\alpha \varepsilon P_k} s\nu_{\mathfrak{D}}(H_\alpha) = \pm \prod_{\alpha \varepsilon P_k} \nu_{\mathfrak{D}}(H_\alpha)$$

and therefore it is obvious that there exists an element $z_0 \varepsilon \mathfrak{U}$ such that

$$z_0(e^{s\nu_{\mathfrak{D}}}) = d(\mathfrak{D})^2 \, | m_{\mathfrak{D}} |^2 \, e^{s\nu_{\mathfrak{D}}} \qquad (s \varepsilon \mathfrak{w}_k).$$

Here $e^{s\nu_{\mathfrak{D}}}$ is the function g on A given by[12] $g(\exp H) = e^{s\nu_{\mathfrak{D}}(H)}$ $(H \varepsilon \mathfrak{h}_0)$. Hence if $g_{\mathfrak{D}}(h) = \eta(h^{-1})\Delta_k(h)\zeta_{\mathfrak{D}}(h)$ it follows that $z_0 g_{\mathfrak{D}} = d(\mathfrak{D})^2 \, | m_{\mathfrak{D}} |^2 g_{\mathfrak{D}}$. Now let ϕ denote the automorphism of $\mathfrak{U}$ over C such that $\phi(H) = -H$

[12] It follows easily from the discussion at the beginning of § 5 that such a function on A actually exists.

and $\phi(1)=1$ $(H\varepsilon\mathfrak{h})$. Then if $z=\phi(z_0)$, $f'(h)=f(h)\eta(h)$ $(h\varepsilon A)$ and m is a positive integer, it is clear that

$$d(\mathfrak{D})^{2m}\,|\,m_{\mathfrak{D}}\,|^{2m}\int_A f'(h)g_{\mathfrak{D}}(h)dh$$
$$=\int_A f'(h)\,(z_0{}^m g_{\mathfrak{D}})\,(h)\,dh=\int_A (z^m f')\,(h)g_{\mathfrak{D}}(h)\,dh.$$

But $|\,e^{s\nu_{\mathfrak{D}}(H)}|=1$ $(H\varepsilon\mathfrak{h}_0)$, hence $|\,g_{\mathfrak{D}}(h)\,|\leqq w$ $(h\varepsilon A)$ where w is the order of $\mathfrak{w}_k$. Hence

$$|\int_A (z^m f')\,(h)g_{\mathfrak{D}}(h)\,dh\,|\leqq w\int_A |\,(z^m f')\,(h)\,|\,dh$$

and therefore

$$\sum_{\mathfrak{D}\varepsilon\Omega_\pi} d(\mathfrak{D})\,|\int_A f(h)\Delta_k(h)\zeta_{\mathfrak{D}}(h)\,dh\,|$$
$$\leqq\sum_{\mathfrak{D}\varepsilon\Omega_\pi} d(\mathfrak{D})^{1-2m}\,|\,m_{\mathfrak{D}}\,|^{-2m}w\int_A |\,z^m f'(h)\,|\,dh.$$

However if m is sufficiently large we know (see [5(c), p. 240]) that

$$\sum_{\mathfrak{D}\varepsilon\Omega_\pi} d(\mathfrak{D})^{1-2m}\,|\,m_{\mathfrak{D}}\,|^{-2m}<\infty$$

and this proves that Q_f is summable.

Now every summable operator has a trace [5(c), Lemma 1]. Let $\tau_\pi(f)$ denote the trace of Q_f. Then the above proof shows that it is possible to choose a positive integer m and a real constant M such that

$$|\,\tau_\pi(f)\,|\leqq M\int_A |\,z^m f'(h)\,|\,dh$$

for all $f\varepsilon C_c{}^\infty(A)$. (Here $f'=f\eta$). This shows that τ_π is a distribution (see Schwartz [8]) on A of finite order. Moreover

$$\tau_\pi(f)=\sum_{\mathfrak{D}\varepsilon\Omega_\pi}\ \sum_{j\varepsilon J(\mathfrak{D})}(\psi_j, Q_f\psi_j)=\sum_{\mathfrak{D}\varepsilon\Omega_\pi} n(\mathfrak{D})\int_A f(h)\Delta_k(h)\zeta_{\mathfrak{D}}(h)\,dh$$

and this shows that τ_π does not change if we replace π by another infinitesimally equivalent representation.

Now let Λ_0 be a real linear function on $\mathfrak{h}$ satisfying the following three conditions (cf. Theorem 3 of [5(f)]):

(1) $\Lambda_0(H_\alpha)$ is a nonnegative integer for every $\alpha\varepsilon P_k$.

(2) $\Lambda_0(H_\beta)=0$ for every noncompact positive root which is not totally positive.

(3) $\Lambda_0(H_\gamma)+\rho(H_\gamma)\leqq 0$ for every totally positive root γ. (Here 2ρ is the sum of all positive roots).

We consider the quasi-simple irreducible representation π of G defined in Theorem 2 corresponding to Λ_0. Put $\tau_{\Lambda_0}=\tau_\pi$ and let $\mathfrak{H}$ denote the representation space of π. Our object is to compute τ_{Λ_0}. We shall see in another paper that τ_{Λ_0} is intimately related with T_{Λ_0} (see however [5(g)]).

We keep to the notation of the proof of Lemma 24. By going over, if necessary, to an equivalent representation we may again assume that π is unitary. For any linear function Λ on $\mathfrak{h}$ let $\mathfrak{H}_\Lambda$ denote the set of all elements $\phi\in\mathfrak{H}$ such that $\pi(\exp H)\phi=e^{\Lambda(H)}\phi$ $(H\in\mathfrak{h}_0)$. Then it is clear that the subspaces $\mathfrak{H}_\Lambda$ are mutually orthogonal. Put $\mathfrak{H}^0=\sum_{\mathfrak{D}\in\Omega}\mathfrak{H}_{\mathfrak{D}}$ and $\mathfrak{H}_\Lambda{}^0=\mathfrak{H}_\Lambda\cap\mathfrak{H}^0$. Since $\mathfrak{H}_{\mathfrak{D}}$ is finite-dimensional and fully reducible under $\pi(A)$, it is obvious that $\mathfrak{H}^0=\sum_\Lambda\mathfrak{H}_\Lambda{}^0$. Let E_Λ denote the orthogonal projection of $\mathfrak{H}$ on $\mathfrak{H}_\Lambda$. Since $\mathfrak{H}^0$ is dense in $\mathfrak{H}$, $E_\Lambda\mathfrak{H}^0$ is dense in $\mathfrak{H}_\Lambda$. But it is clear that $E_\Lambda\mathfrak{H}^0\subset\mathfrak{H}^0$ and therefore $E_\Lambda\mathfrak{H}^0\subset\mathfrak{H}_\Lambda{}^0$. However $\dim\mathfrak{H}_\Lambda{}^0<\infty$ (corollary to Lemma 21 of [5(f)]) and so it follows that $\mathfrak{H}_\Lambda=\mathfrak{H}_\Lambda{}^0$.

Let P_+ denote the set of all totally positive roots of $\mathfrak{g}$ and let $\mathfrak{h}_+$ be the open subset of $\mathfrak{h}$ consisting of all those element H for which $|e^{\gamma(H)}|>1$ for every $\gamma\in P_+$. Since every root takes pure imaginary values on $\mathfrak{h}_0$, it is obvious that $\mathfrak{h}_++\mathfrak{h}_0\subset\mathfrak{h}$. Let σ_β $(\beta\in P_+)$ denote the hyperplane in the real Euclidean space $\mathfrak{h}^*=(-1)^{\frac{1}{2}}\mathfrak{h}_0$ defined by the equation $\beta(H)=0$. Consider the complement $\mathfrak{h}_1{}^*$ of $\bigcup_{\beta\in P_+}\sigma_\beta$ in $\mathfrak{h}^*$. Let $(\alpha_1,\cdots,\alpha_l)$ be a fundamental system of positive roots and H_0 a point in $\mathfrak{h}$ such that $\alpha_i(H_0)=1$ $1\leqq i\leqq l$. Then $H_0\in\mathfrak{h}_1{}^*$ and if $\mathfrak{h}_+{}^*$ is the connected component of H_0 in $\mathfrak{h}_1{}^*$, it is obvious that $\mathfrak{h}_+{}^*\subset\mathfrak{h}^*\cap\mathfrak{h}_+$. Conversely $\mathfrak{h}^*\cap\mathfrak{h}_+$ is a convex subset of $\mathfrak{h}_1{}^*$ and therefore it is connected. Hence $\mathfrak{h}_+{}^*=\mathfrak{h}^*\cap\mathfrak{h}_+$ and therefore $\mathfrak{h}_+=\mathfrak{h}_0+\mathfrak{h}_+{}^*$. In particular $tH_0\in\mathfrak{h}_+{}^*$ for every $t>0$ and so zero lies in the closure of $\mathfrak{h}_+$. Now consider the complex abelian Lie group $\tilde{A}_c$ defined in Section 3 and let ${}_+\tilde{A}_c$ denote the subset of those $h\in\tilde{A}_c$ which can be written in the form $h=\exp H$ $(H\in\mathfrak{h}_+)$. It is clear that ${}_+\tilde{A}_c$ is an open connected subset of $\tilde{A}_c$ whose closure contains 1. Also ${}_+\tilde{A}_cA={}_+\tilde{A}_c$. We shall now define a bounded operator $\pi(h)$ on $\mathfrak{H}$ for every $h\in{}_+\tilde{A}_c$. Let $\mathfrak{F}_\pi$ be the set of those linear function Λ on $\mathfrak{h}$ for which $\mathfrak{H}_\Lambda\neq\{0\}$. Then if $\exp H=1$ in A $(H\in\mathfrak{h}_0)$ it is obvious that $e^{\Lambda(H)}=1$ and therefore, as we have seen in Section 5, there exists a holomorphic character ξ_Λ of $\tilde{A}_c$ such that $\xi_\Lambda(\exp H)=e^{\Lambda(H)}$ $(H\in\mathfrak{h})$. Let ψ be the element in $\mathfrak{H}$ corresponding to Theorem 2. Then $\psi\in\mathfrak{H}_{\Lambda_0}$ and $\mathfrak{H}^0=\pi(\mathfrak{B})\psi$. Moreover if $V=\pi(\mathfrak{X})\psi$, $\dim V<\infty$ (Lemmas 8 and 10) and

V is irreducible under $\pi(\mathfrak{k})$ (Lemma 2 of [5(f)]). Let $\mathfrak{D}_0$ denote the class[13] (in Ω_π) and $\Lambda_0, \Lambda_1, \cdots, \Lambda_r$ all the (distinct) weights of this representation of $\mathfrak{k}$ on V. Then if $\gamma_1, \cdots, \gamma_q$ are all the (distinct) totally positive roots of $\mathfrak{g}$, every $\Lambda \varepsilon \mathfrak{F}_\pi$ can be written in the form

$$\Lambda = \Lambda_i - (m_1\gamma_1 + \cdots + m_q\gamma_q)$$

for some i $(0 \leqq i \leqq r)$ and some non-negative integers $m_1, \cdots, m_q$ (see the corollary to Lemma 21 of [5(f)]). Therefore if $h \varepsilon {}_+\tilde{A}_c$

$$|\xi_\Lambda(h)| \leqq \max_{0 \leqq i \leqq r} |\xi_{\Lambda_i}(h)| \qquad (\Lambda \varepsilon \mathfrak{F}_\pi).$$

Hence it is clear that the infinite sum $\sum_{\Lambda \varepsilon \mathfrak{F}_\pi} \xi_\Lambda(h) E_\Lambda \phi$ converges in $\mathfrak{H}$ for any $h \varepsilon {}_+\tilde{A}_c$ and $\phi \varepsilon \mathfrak{H}$. Let $\pi(h)\phi$ denote the limit of this sum. Then

$$|\pi(h)\phi|^2 \leqq \max_{0 \leqq i \leqq r} |\xi_{\Lambda_i}(h)|^2 |\phi|^2$$

and therefore $\pi(h)$ is a bounded operator on $\mathfrak{H}$. It is obvious from its definition that if $h_+ \varepsilon {}_+\tilde{A}_c$ and $h \varepsilon A$, $\pi(hh_+) = \pi(h)\pi(h_+) = \pi(h_+)\pi(h)$.

Let $\zeta_{\mathfrak{D}_0}$ denote the character of the class $\mathfrak{D}_0$. Then if $d_i = \dim(V \cap \mathfrak{H}_{\Lambda_i})$ $0 \leqq i \leqq r$,

$$\zeta_{\mathfrak{D}_0}(h) = \sum_{0 \leqq i \leqq r} d_i \xi_{\Lambda_i}(h)$$

for any $h \varepsilon A$ and we can extend $\zeta_{\mathfrak{D}_0}$ to a holomorphic function on $\tilde{A}_c$ by means of the above formula. Let λ be a linear function on $\mathfrak{h}$ such that $\lambda(H_\alpha)$ is an integer for every compact root α. Then as we have seen in Section 5, there exists a holomorphic character ξ_λ of $\tilde{A}_c$, such that $\xi_\lambda(\exp H) = e^{\lambda(H)}$ $(H \varepsilon \mathfrak{h})$. Put $\rho_+ = \frac{1}{2} \sum_{\gamma \varepsilon P_+} \gamma$, $\rho_k = \frac{1}{2} \sum_{\alpha \varepsilon P_k} \alpha$ and $\rho_0 = \rho_+ + \rho_k$. Then if s_α is the Weyl reflexion corresponding to a compact root α, it follows from Lemma 10 of [5(f)] that $s_\alpha \rho_+ = \rho_+$ and therefore $\rho_+(H_\alpha) = 0$. Also if β is any root, $\beta(H_\alpha)$ and $\rho_k(H_\alpha)$ are integers (see Weyl [9]). This shows that we can construct the corresponding characters ξ_{ρ_+}, ξ_{ρ_k}, ξ_{ρ_0}, ξ_β. Put

$$\Delta_k(h) = \xi_{\rho_k}(h) \prod_{\alpha \varepsilon P_k} \{1 - \xi_\alpha(h^{-1})\} = \prod_{\alpha \varepsilon P_k} (e^{\frac{1}{2}\alpha(H)} - e^{-\frac{1}{2}\alpha(H)}),$$

$$\Delta_+(h) = \xi_{\rho_+}(h) \prod_{\gamma \varepsilon P_+} \{1 - \xi_\gamma(h^{-1})\} = \prod_{\gamma \varepsilon P_+} \{e^{\frac{1}{2}\gamma(H)} - e^{-\frac{1}{2}\gamma(H)}\}$$

where $h = \exp H \varepsilon \tilde{A}_c$ $(H \varepsilon \mathfrak{h})$.

LEMMA 25. *If $h \varepsilon {}_+\tilde{A}_c$ the operator $\pi(h)$ is summable and*

$$sp\,\pi(h) = \{\Delta_+(h)\}^{-1} \xi_{\rho_+}(h) \zeta_{\mathfrak{D}_0}(h).$$

---

[13] As usual we identify finite-dimensional irreducible representations of $\mathfrak{k}$ with those of K.

Moreover the series $\sum_{\Lambda \varepsilon \mathfrak{F}_\pi} (\dim \mathfrak{H}_\Lambda) |\xi_\Lambda(h)|$ *converges uniformly on every compact subset of* ${}_+\tilde{A}_c$.

Let $h = \exp H$ $(H \varepsilon \mathfrak{h}_+)$. Then from the corollary to Lemma 21 of [5(f)],

$$\sum_{\Lambda \varepsilon \mathfrak{F}_\pi} (\dim \mathfrak{H}_\Lambda) |\xi_\Lambda(h)|$$

$$= \sum_{0 \leqq i \leqq r} d_i \sum_{m_1, \cdots, m_q \geqq 0} |\exp\{\Lambda_i(H) - m_1\gamma_1(H) - \cdots - m_q\gamma_q(H)\}|.$$

Since the real part of $\gamma_j(H)$ is positive, the series converges and clearly its convergence is uniform if H remains in a compact subset of $\mathfrak{h}_+$. Moreover since $\pi(h)$ coincides with $\xi_\Lambda(h)E_\Lambda$ on $\mathfrak{H}_\Lambda$, it is now clear that $\pi(h)$ is summable and

$$sp\pi(h) = \sum_{\Lambda \varepsilon \mathfrak{F}_\pi} (\dim \mathfrak{H}_\Lambda)\xi_\Lambda(h)$$

$$= \sum_{0 \leqq i \leqq r} d_i \sum_{m_1, \cdots, m_q \geqq 0} \exp(\Lambda_i(H) - m_1\gamma_1(H) - \cdots - m_q\gamma_q(H))$$

$$= \prod_{1 \leqq j \leqq q} \{1 - e^{-\gamma_j(H)}\}^{-1} \zeta_{\mathfrak{D}_0}(\exp H) = \{\Delta_+(h)\}^{-1} \xi_{\rho_+}(h) \zeta_{\mathfrak{D}_0}(h).$$

For any $f \varepsilon C_c^\infty(A)$ and $h_+ \varepsilon {}_+\tilde{A}_c$, put

$$Q_f(h_+) = \int_A f(h)\Delta_k(hh_+)\pi(hh_+)dh, \quad Q_f = \int_A \int_{K_0} f(h)\Delta_k(h)\pi(h^{\bar{u}})dhd\bar{u}.$$

LEMMA 26. *If* $f \varepsilon C_c^\infty(A)$ *the operator* $Q_f(h_+)$ *is summable and*

$$\lim_{h_+ \to 1} spQ_f(h_+) = spQ_f \qquad (h_+ \varepsilon {}_+\tilde{A}_c).$$

Let $\Lambda \varepsilon \mathfrak{F}_\pi$. Then it is obvious that

$$Q_f(h_+)E_\Lambda = \{\int_A f(h)\Delta_k(hh_+)\xi_\Lambda(hh_+)dh\}E_\Lambda,$$

and therefore, in order to show that $Q_f(h_+)$ is summable, it is enough to prove that

$$\sum_{\Lambda \varepsilon \mathfrak{F}_\pi} (\dim \mathfrak{H}_\Lambda) |\int_A f(h)\Delta_k(hh_+)\xi_\Lambda(hh_+)dh| < \infty.$$

But this is obvious in view of the fact that

$$\sum_{\Lambda \varepsilon \mathfrak{F}_\pi} (\dim \mathfrak{H}_\Lambda) |\xi_\Lambda(hh_+)|$$

converges uniformly (with respect to h) on every compact subset of A (Lemma 25).

Now we use the notation of the proof of Lemma 24. Since $Q_f(h_+)$ is summable,

$$spQ_f(h_+) = \sum_{\mathfrak{D}\,\varepsilon\,\Omega_\pi} sp(E_{\mathfrak{D}}Q_f(h_+)E_{\mathfrak{D}}).$$

Then if we define η, μ, $\nu_{\mathfrak{D}}$, $n(\mathfrak{D})$ and $m(\mathfrak{D})$ $(\mathfrak{D}\,\varepsilon\,\Omega_\pi)$ as before,

$$\eta(\exp H) = e^{\mu(H)}, \Delta_k(\exp H)\zeta_{\mathfrak{D}}(\exp H) = e^{\mu(H)} \sum_{s\,\varepsilon\,\mathfrak{w}_k} \epsilon(s) e^{s\nu_{\mathfrak{D}}(H)} \qquad (H\,\varepsilon\,\mathfrak{h}_0)$$

and it is clear that we can extend η and $\zeta_{\mathfrak{D}}$ to holomorphic functions on $\tilde{A}_c$ so that the above relations actually hold for all $H\,\varepsilon\,\mathfrak{h}$. Then if $\mathfrak{D}\,\varepsilon\,\Omega_\pi$,

$$sp(E_{\mathfrak{D}}Q_f(h_+)E_{\mathfrak{D}}) = n(\mathfrak{D}) \int_A f(h)\Delta_k(hh_+)\zeta_{\mathfrak{D}}(hh_+)\,dh$$

and therefore

$$spQ_f(h_+) = \sum_{\mathfrak{D}\,\varepsilon\,\Omega_\pi} n(\mathfrak{D}) \int_A f(h)\Delta_k(hh_+)\zeta_{\mathfrak{D}}(hh_+)\,dh.$$

Now put $\nu'_{\mathfrak{D}} = \nu_{\mathfrak{D}} + \mu_0$ where μ_0 is the restriction of μ on $\mathfrak{h}$,

$$g_{\mathfrak{D}}(h) = \eta(h^{-1})\Delta_k(h)\zeta_{\mathfrak{D}}(h) = \sum_{s\,\varepsilon\,\mathfrak{w}_k} \epsilon(s)\xi_{s\nu_{\mathfrak{D}}}(h) \qquad (h\,\varepsilon\,\tilde{A}_c)$$

and $f'(h) = f(h)\eta(h)$ $(h\,\varepsilon\,A)$. It is obvious from the definition of $\nu'_{\mathfrak{D}}$ that $\nu'_{\mathfrak{D}} - \rho_k\,\varepsilon\,\mathfrak{F}_\pi$ and therefore $s(\nu'_{\mathfrak{D}} - \rho_k)\,\varepsilon\,\mathfrak{F}_\pi$ for $s\,\varepsilon\,\mathfrak{w}_k$ (corollary to Lemma 6 of [5(f)]). Hence

$$|\xi_{s\nu_{\mathfrak{D}}}(h)| \leqq \max_{\substack{0\leqq i\leqq r\\ s\,\varepsilon\,\mathfrak{w}_k}} |\xi_{\Lambda_i+s\rho_k}(h)|\ |\eta(h^{-1})| \qquad (h\,\varepsilon\,{}_+\tilde{A}_c).$$

We can now argue as in the proof of Lemma 24 and show that the series

$$\sum_{\mathfrak{D}\,\varepsilon\,\Omega_\pi} n(\mathfrak{D}) \left| \int_A f(h)\Delta_k(hh_+)\zeta_{\mathfrak{D}}(hh_+)\,dh \right|$$

converges uniformly with respect to h_+ on $B \cap {}_+\tilde{A}_c$ where B is a compact neighbourhood of 1 in $\tilde{A}_c$. Hence it follows that

$$\begin{aligned}\lim_{h_+\to 1} spQ_f(h_+) &= \sum_{\mathfrak{D}\,\varepsilon\,\Omega_\pi} n(\mathfrak{D}) \lim_{h_+\to 1} \int_A f(h)\Delta_k(hh_+)\zeta_{\mathfrak{D}}(hh_+)\,dh \\ &= \sum_{\mathfrak{D}\,\varepsilon\,\Omega_\pi} n(\mathfrak{D}) \int_A f(h)\Delta_k(h)\zeta_{\mathfrak{D}}(h)\,dh = spQ_f.\end{aligned}$$

Now put

$$\Xi_{\Lambda_0} = \xi_{\rho_+}\zeta_{\mathfrak{D}_0}\Delta_k = \sum_{s\,\varepsilon\,\mathfrak{w}_k} \epsilon(s)\xi_{s(\Lambda_0+\rho_0)}.$$

COROLLARY 1. $\tau_{\Lambda_0}(f) = \operatorname{Lim}_{h_+\to 1} \int_A f(h)\,(\Xi_{\Lambda_0}(hh_+)/\Delta_+(hh_+))\,dh \quad (h_+ \varepsilon\, {}_+\tilde{A}_c)$ *for* $f \varepsilon C_c^\infty(A)$.

This is an immediate consequence of Lemma 26 and the fact that $\tau_{\Lambda_0}(f) = spQ_f$.

COROLLARY[14] 2. $\Delta_+\tau_{\Lambda_0} = \Xi_{\Lambda_0}$ *on* A.

Let H_+ be an element in $\mathfrak{h}_+^*$ so that $\gamma(H_+)$ is real and positive for every $\gamma \varepsilon P_+$. Then if $h_+(t) = \exp tH_+$ $(t > 0)$ we know from the above lemma that

$$\tau_{\Lambda_0}(\Delta_+ f) = \operatorname{Lim}_{t\to 0} \int_A \Delta_+(h) f(h) \{\Xi_{\Lambda_0}(hh_+(t))/\Delta_+(hh_+(t))\} dh$$

for $f \varepsilon C_c^\infty(A)$. Now we claim that there exists a constant M such that

$$|\Delta_+(h)\Xi_{\Lambda_0}(hh_+(t))/\Delta_+(hh_-(t))| \leqq M$$

for all $h \varepsilon A$ provided t is sufficiently small and positive. This is seen as follows. If θ and ϵ are real numbers and $\epsilon \geqq \frac{1}{2}$,

$$\begin{aligned} |(e^{(-1)^{\frac{1}{2}}\theta} - 1)/(\epsilon e^{(-1)^{\frac{1}{2}}\theta} - 1)|^2 &= (2 - 2\cos\theta)/(1 + \epsilon^2 - 2\epsilon\cos\theta) \\ &= 2(1-\cos\theta)/\{(1-\epsilon)^2 + 2\epsilon(1-\cos\theta)\} \leqq 1/\epsilon \leqq 2. \end{aligned}$$

Hence if q is the number of totally positive roots of $\mathfrak{g}$, it is clear that

$$|\Delta_+(h)/\Delta_+(hh_+(t))| \leqq 2^q$$

provided t is sufficiently small and positive. Our assertion is now obvious. Therefore by Lebesgue's Theorem

$$\tau_{\Lambda_0}(\Delta_+ f) = \int_A f(h) \operatorname{Lim}_{t\to 0} \{\Delta_+(h)\Xi_{\Lambda_0}(hh_+(t))/\Delta_+(hh_+(t))\} dh.$$

But

$$\operatorname{Lim}_{t\to 0} \Delta_+(h)\Xi_{\Lambda_0}(hh_+(t))/\Delta_+(hh_+(t)) = \begin{cases} 0 \text{ if } \Delta_+(h) = 0 \\ \Xi_{\Lambda_0}(h) \text{ if } \Delta_+(h) \neq 0 \end{cases} \quad (h \varepsilon A)$$

and since the set of all $h \varepsilon A$ for which $\Delta_+(h) = 0$ is of Haar measure zero, we get

$$\tau_{\Lambda_0}(\Delta_+ f) = \int_A f(h)\Xi_{\Lambda_0}(h)\,dh.$$

This proves the corollary.

COLUMBIA UNIVERSITY.

[14] Here we use the standard terminology of the theory of distributions (Schwartz [8]).

REFERENCES.

[1] G. Birkhoff, "Representability of Lie algebras and Lie groups by matrices," *Annals of Mathematics*, vol. 38 (1937), pp. 526-532.

[2] S. Bochner and W. T. Martin, *Several Complex Variables*, Princeton University Press, 1948.

[3] E. Cartan, "Sur certaines formes riemanniennes remarquables des géométries à groupe fondamental simple," *Annales Scientifiques de l'École Normale Supérieure*, vol. 44 (1927), pp. 345-467.

[4] C. Chevalley, *Theory of Lie Groups*, Princeton University Press, 1946.

[5] Harish-Chandra, (a) "On some applications of the universal enveloping algebra of a semisimple Lie algebra," *Transactions of the American Mathematical Society*, vol. 70 (1951), pp. 28-96.

———, (b) "Representations of a semisimple Lie group on a Banach space. I," *ibid.*, vol. 75 (1953), pp. 185-243.

———, (c) "Representations of semisimple Lie groups. III," *ibid.*, vol. 76 (1954), pp. 234-253.

———, (d) "The Plancherel formula for complex semisimple Lie groups," *ibid.*, vol. 76 (1954), pp. 485-528.

———, (e) "Integrable and square-integrable representations of a semisimple Lie group," *Proceedings of the National Academy of Sciences*, vol. 41 (1955), pp. 314-317.

———, (f) "Representations of semisimple Lie groups. IV," *American Journal of Mathematics* vol. 77 (1955) pp. 743-777.

———, (g) "On the characters of a semisimple Lie group," *Bulletin of the American Mathematical Society*, vol. 61 (1955), pp. 389-396.

[6] K. Iwasawa, "On some types of topological groups," *Annals of Mathematics*, vol. 50 (1949), pp. 507-557.

[7] G. D. Mostow, "A new proof of E. Cartan's theorem on the topology of semi-simple groups," *Bulletin of the American Mathematical Society*, vol. 55 (1949), pp. 969-980.

[8] L. Schwartz, *Théorie des distributions*, Paris, Hermann, 1950.

[9] H. Weyl, *The Structure and Representation of Continuous Groups*, The Institute for Advanced Study, Princeton, 1935.

SPECTRAL ISOMORPHISMS FOR SOME RINGS OF INFINITE MATRICES ON A BANACH SPACE.*

By G. L. KRABBE.

1. **Introduction.** Suppose $p > 1$, and let l_p denote the set of all sequences c such that $\sum_{n=-\infty}^{\infty} | c_n |^p < \infty$. If a is a sequence $\{a_n\}$, the matrix $(a_{n-\nu})$ is usually known as a *Laurent matrix* when $n, \nu = 0, \pm 1, \pm 2, \cdots$; it represents the transformation $\langle a \rangle_p$ which maps any member c of l_p on a sequence b defined by

$$b_n = \sum_{\nu=-\infty}^{\infty} a_{n-\nu} c_\nu \qquad (n = 0, \pm 1, \pm 2, \cdots).$$

If F belongs to the ring (L^∞) of essentially bounded summable functions on $[-\pi, \pi]$, then ΛF will denote the sequence of Fourier coefficients of F. Restricting themselves to the case $p = 2$, O. Toeplitz [14, pp. 499-502] and F. Riesz [12] have shown that, when $\mathfrak{F} = (L^\infty)$, then

(i) *the mapping $F \to \langle \Lambda F \rangle_p$ is a continuous isomorphism of $\mathfrak{F}$ into the ring $\mathfrak{E}$ of bounded operators on l_p.*

When $\mathfrak{F}$ is the ring (C) of continuous functions (and again with the restriction $p = 2$), they established (see [4]) that

(ii) *the spectrum of $\langle \Lambda F \rangle_p$ is the image $F([-\pi, \pi])$, when $F \varepsilon \mathfrak{F}$,*

(iii) *if $F \varepsilon \mathfrak{F}$, then the inverse Q of the operator $\langle \Lambda F \rangle_p$ exists if and only if F does not vanish on $[-\pi, \pi]$. When Q exists, then $Q = \langle \Lambda G \rangle_p$, where $G(\theta) = [F(\theta)]^{-1}$.*

Assume henceforth $p > 1$; we will show in 4.2 that (i) holds when $\mathfrak{F}$ is the ring (BV) of functions of bounded variation on $[-\pi, \pi]$. The above results of Riesz and Toeplitz prompted us to consider (in [6]) the extent to which (ii) holds when F is in the ring (CBV) of continuous functions in (BV); it was found that the spectrum of $\langle \Lambda F \rangle_p$ is then a connected subset of $F([-\pi, \pi])$. In the present paper, we prove that (ii) and (iii) hold

* Received June 30, 1955; revised October 27, 1955.

when $\mathfrak{F}$ is the ring (AC) of absolutely continuous functions. From these results follows that, if $F \varepsilon (AC)$ and $T = \langle \Delta F \rangle_p$, then

$$\sup\{|F(\theta)| : |\theta| \leqq \pi\} \leqq \| T \|.$$

The equality sign holds when $p = 2$ (see 4.4 and Riesz [12]). In the general case $p > 1$, we show that T belongs to the ring $\mathfrak{E}[H]$ generated by the operator H which is represented by the Laurent matrix $(\alpha_{n\nu})$, where $\alpha_{n\nu} = i(-1)^{n+\nu}/(n-\nu)$ and $\alpha_{nn} = 0$. Moreover, the mapping $F \to \langle \Delta F \rangle_p$ is the only continuous homomorphism of (AC) into $\mathfrak{E}$ which maps on H the identity-function.

1. 1. Application. Suppose f belongs to the ring $\mathfrak{U}$ of all functions f such that $f(\lambda) = \sum a_n \lambda^n$ $(n = -\infty \cdots \infty)$ for some a in l_1. Note that the perceding series is the Laurent-expansion of f in some annulus containing $\Gamma_1 = \{\lambda : |\lambda| = 1\}$; two members of $\mathfrak{U}$ will be considered equal when they coincide on Γ_1. Set $\Delta f = \Delta F$, where $F(\theta) = f(e^{i\theta})$; since $F(\theta) = \sum a_n e^{in\theta}$, we have $a = \Delta f$. The Laurent matrix $(a_{n-\nu})$ again represents an operator $\langle \Delta f \rangle_p$. Toeplitz ([15], [16]) has proved that, when $p = 2$ and $\mathfrak{F} = \mathfrak{U}$, then

(i′) *the mapping $f \to \langle \Delta f \rangle_p$ is an isomorphism of $\mathfrak{F}$ into the ring $\mathfrak{E}$,*

(ii′) *the spectrum of $\langle \Delta f \rangle_p$ is the image of Γ_1 by f, when $f \varepsilon \mathfrak{F}$,*

(iii′) *if $f \varepsilon \mathfrak{F}$, then the inverse Q of the operator $\langle \Delta f \rangle_p$ exists if and only if f does not vanish on Γ_1. When Q exists, then $Q = \langle \Delta g \rangle_p$, where $g(\lambda) = [f(\lambda)]^{-1}$.*

It follows readily from our results that (i′)-(iii′) hold for any $p > 1$, when we take for $\mathfrak{F}$ the larger class of all functions f such that $f(e^{i\theta})$ is an absolutely continuous function of θ $(|\theta| \leqq \pi)$.

2. Spectral mappings. Suppose $\mathfrak{R}$ is a Banach algebra with unit $\mathbf{1}$; we write $R = \mathfrak{R}\,\mathrm{lm}\, R_n$ if $\lim \| R - R_n \| = 0$ ($n \to \infty$, and in the norm of $\mathfrak{R}$). Let $\mathfrak{R}'$ be the set of all T in $\mathfrak{R}$ such that T has an inverse (denoted T^{-1}) in $\mathfrak{R}$. The spectrum $\sigma(R)$ of some R in $\mathfrak{R}$ is the set of all complex λ such that $\lambda\mathbf{1} - R \notin \mathfrak{R}'$.

Suppose $\mathfrak{S}$ is a fixed compact subset of the plane. The Banach algebra $\boldsymbol{C}(\mathfrak{S})$ of all complex-valued functions on $\mathfrak{S}$ has the norm

$$\| A \|_\infty = \sup\{|A(\theta)| : \theta \varepsilon \mathfrak{S}\};$$

the product $A \cdot B$ of two members of $\boldsymbol{C}(\mathfrak{S})$ is the function F defined by $F(\theta) = A(\theta) \cdot B(\theta)$.

2.1 *Remark.* Denoting by I the member of $\boldsymbol{C}(\mathfrak{s})$ defined by $I(\theta)=\theta$, we have $I^n(\theta)=\theta^n$ and $I^0(\theta)=1$; note that I^0 is the unit-element of $\boldsymbol{C}(\mathfrak{s})$. It is easily verified that $\sigma(A)$ is the image $A(\mathfrak{s})$ of $\mathfrak{s}$ by A (when $A \varepsilon \boldsymbol{C}(\mathfrak{s})$).

2.2 DEFINITION. *Assume that $\mathfrak{T}$ is a subset of a given Banach algebra. If ϕ is a mapping of $\mathfrak{T}$ into some Banach algebra, we say that ϕ is "spectral" if $\sigma(T)=\sigma(\phi(T))$ for all T in $\mathfrak{T}$.*

2.3 LEMMA. *If $\mathfrak{T}$ is a subset of some Banach algebra, and if ϕ is a spectral mapping of $\mathfrak{T}$ into $\boldsymbol{C}(\mathfrak{s})$, then $\|\phi(T)\| \leqq \|T\|$ when $T \varepsilon \mathfrak{T}$.*

Proof. Set $A=\phi(T)$. It is easily seen that $\|A\|_\infty=\sup\{|\lambda|:\lambda\varepsilon\sigma(A)\}$, and therefore $\|A\|_\infty=\sup\{|\lambda|:\lambda\varepsilon\sigma(T)\}\leqq\lim\|T^n\|^{1/n}\leqq\|T\|$; the first inequality always holds for a member T of some Banach algebra ([7], 24 A).

2.4 LEMMA. *Let $\mathfrak{F}$ denote either $\boldsymbol{C}(\mathfrak{s})$ or one of its subsets (AC) and (CBV). Suppose Φ is a spectral homomorphism of some Banach algebra $\mathfrak{X}$ into $\boldsymbol{C}(s)$. If $T\varepsilon\mathfrak{X}$ and $A=\Phi(T)\varepsilon\mathfrak{F}$, then*

(iv) $\|A\|_\infty\leqq\|T\|$; $\sigma(T)=A(\mathfrak{s})$; $T^{-1}\varepsilon\mathfrak{X}\leftrightarrows 0\notin A(\mathfrak{s})$,

(v) *if* $T^{-1}\varepsilon\mathfrak{X}$ *then* $A^{-1}=\Phi(T^{-1})\varepsilon\mathfrak{F}$.

Remark. Since $\sigma(A)=A(\mathfrak{s})$ (see 2.1), we have $0\notin A(\mathfrak{s})\leftrightarrows A(\theta)\neq 0$ for all θ in $\mathfrak{s}\leftrightarrows 0\notin\sigma(A)$; note that $A^{-1}(\theta)=[A(\theta)]^{-1}$ when $0\notin A(\mathfrak{s})$.

Proof. The first two statements of (iv) follow from 2.3, 2.2, and 2.1. On the other hand, $T^{-1}\varepsilon\mathfrak{X}\leftrightarrows 0\notin\sigma(T)=\sigma(A)\leftrightarrows A^{-1}\varepsilon\mathfrak{F}$; the proof is now concluded by noting that $A^{-1}\varepsilon\mathfrak{F}\leftrightarrows 0\notin A(\mathfrak{s})$ in each of the cases considered.

2.5 *Remark.* If $\mathfrak{s}$ is connected and if $\mathfrak{X}$ is abelian, then the existence of a spectral mapping of $\mathfrak{X}$ into $\boldsymbol{C}(\mathfrak{s})$ implies that $\mathfrak{X}$ is irreducible (in the sense that $\mathbf{1}$ and $\mathbf{O}$ are the only members X of $\mathfrak{X}$ such that $X^2=X$). This is because a necessary and sufficient condition for irreducibility is that the spectra of all members be connected sets ([5], p. 454).

2.6 *Remark.* The inverse V of the mapping $F\to\Lambda F$ is a spectral isomorphism of the Banach algebra l_1 into $\boldsymbol{C}(\mathfrak{s})$ (where $\mathfrak{s}=[-\pi,\pi]$; see 4.5). A related example of a spectral isomorphism is also found in 4.5. Note that (iv)-(v) include the statements (ii)-(iii) when $T=\langle\Lambda A\rangle_p$.

2.7 THEOREM. *Suppose $\mathfrak{T}$ is dense in an abelian Banach algebra $\mathfrak{X}$. If ϕ is a spectral homomorphism of $\mathfrak{T}$ into $\boldsymbol{C}(\mathfrak{s})$, then ϕ can be extended to a spectral homomorphism Φ of $\mathfrak{X}$ into $\boldsymbol{C}(\mathfrak{s})$.*

Proof. The following property will be needed. If R and R_n are members of some abelian Banach algebra $\mathfrak{R}$, then (assuming $n \to \infty$ throughout)

(a) $$R = \mathfrak{R} \operatorname{lm} R_n \quad \text{implies} \quad \sigma(R) = \lim \sigma(R_n).$$

In this statement (proved in [11]), lim is the Hausdorff limit. From the continuity of ϕ (see 2.3) follows (using a well known theorem; cf. [7], 7 F) that ϕ can be extended to a continuous mapping Φ on $\mathfrak{X}$;

(1) $$\Phi(T) = \phi(T) \text{ for all } T \text{ in } \mathfrak{T}.$$

The proof is completed by establishing the spectrality of Φ. For any X in $\mathfrak{X}$ there exists some T_n in $\mathfrak{T}$ such that $X = \mathfrak{X} \operatorname{lm} T_n$; accordingly, if Ψ is a continuous mapping on $\mathfrak{X}$, then (by (a)),

(2) $$\sigma(\Psi(X)) = \lim \sigma(\Psi(T_n)) \qquad (n \to \infty).$$

This holds therefore for $\Psi = \Phi$ and $\Psi = I$ (the identity mapping $I(R) = R$). On the other hand, $\sigma(\Phi(T_n)) = \sigma(\phi(T_n)) = \sigma(I(T_n))$ follows from (1) and the spectrality of ϕ. The conclusion $\sigma(\Phi(X)) = \sigma(I(X)) = \sigma(X)$ is now a consequence of (2).

2.8 Definitions. *The set l_0 consists of all sequences a such that $a_n = 0$ except for finitely many non-negative values of n. We say that $P \varepsilon \mathfrak{P}$ if $P = \sum a_n I^n$ for some a in l_0 (recall that $I^n(\theta) = \theta^n$); clearly $a_n = P^{(n)}(0)/n!$ and $\mathfrak{P}$ is the family of all polynomials. In case J belongs to some Banach algebra $\mathfrak{E}$, we denote by $[J; \mathfrak{E}]$ the mapping $P \to \sum (P^{(n)}(0)/n!) J^n$ of $\mathfrak{P}$ into $\mathfrak{E}$; the closure in $\mathfrak{E}$ of the range of $[J; \mathfrak{E}]$ is the subalgebra $\mathfrak{E}[J]$ generated by J.*

2.9 Lemma. *Suppose J is a member of a Banach algebra $\mathfrak{E}$ such that $\sigma(J)$ is an infinite set $\mathfrak{S}$. If $\psi = [J; \mathfrak{E}]$, then ψ is an isomorphism of $\mathfrak{P}$ into $\mathfrak{E}$, and $\psi(I) = J$. Any homomorphism ψ_1 of $\mathfrak{P}$ into $\mathfrak{E}$ which maps I on J is identical to ψ.*

Proof. If $P \varepsilon \mathfrak{P}$, then $P = \sum a_n I^n$ and $\psi(P) = \sum a_n J^n$ for some a in l_0. By the Dunford mapping theorem ([3], 2.8 and 2.9)

(3) $$\sigma(\psi(P)) = P(\sigma(J)) = P(\mathfrak{S}) \qquad (\text{when } P \varepsilon \mathfrak{P}).$$

It is easily checked that ψ is a homomorphism. Suppose $\psi(P) = \mathrm{O}$; this implies (taking $\{0\} = \sigma(\mathrm{O})$ and (3) into account) $\{0\} = \sigma(\psi(P)) = P(\mathfrak{S})$. Hence the polynomial P vanishes at all points of the infinite (and bounded) set $\mathfrak{S}$; consequently $P = \mathrm{O}$. This proves that the linear transformation ψ

is an isomorphism. We conclude by observing that the homomorphism ψ_1 satisfies $\psi_1(\sum a_n I^n) = \sum a_n[\psi_1(I)]^n$, so that $\psi_1 = \psi$ when $\psi_1(I) = J$.

2.10 LEMMA. *Let $\mathfrak{E}$, J, and $\mathfrak{S}$ be as in 2.9. There exists a spectral homomorphism Φ of $\mathfrak{E}[J]$ into $\boldsymbol{C}(\mathfrak{S})$ which maps J on I. If $\psi = [J;\mathfrak{E}]$, then $P = \Phi(\psi(P))$ when $P \varepsilon \mathfrak{P}$.*

Proof. If $P \varepsilon \mathfrak{P}$ and $T = \psi(P)$, then $\sigma(T) = P(\mathfrak{S}) = \sigma(P)$ (by (3), $P \varepsilon \boldsymbol{C}(\mathfrak{S})$, and 2.1). Hence the inverse ϕ of ψ is a spectral homomorphism of the ring $\psi(\mathfrak{P}) = \mathfrak{T}$ into $\boldsymbol{C}(\mathfrak{S})$. Since $\mathfrak{T}$ is dense in the Banach algebra $\mathfrak{E}[J]$, the conclusion now follows from 2.7.

2.11 *Remark.* From 2.3 follows that the mapping Φ of 2.10 is a continuous homomorphism of $\mathfrak{E}[J]$ into $\boldsymbol{C}(\mathfrak{S})$ mapping J on I. We will now show that it is the only such mapping. Take $\iota = 1, 2$ and let Φ_ι be continuous homomorphisms of $\mathfrak{E}[J]$ into $\boldsymbol{C}(\mathfrak{S})$ such that $\Phi_\iota(J) = I$: note that Φ_1 and Φ_2 coincide on the set $\psi(\mathfrak{P}) = \{\sum a_n J^n : a \varepsilon l_0\}$ (since $\Phi_\iota(\sum a_n J^n) = \sum a_n[\Phi_\iota(J)]^n = \sum a_n I^n$); but Φ_1 and Φ_2, being also continuous on the closure $\mathfrak{E}[J]$ of $\psi(\mathfrak{P})$, must consequently coincide on $\mathfrak{E}[J]$ (cf. [7]; 7 F). The existence of such a mapping could equally well have been obtained by appealing to Gelfand's theory of maximal ideals ([7], 23 E).

2.12 DEFINITION. *Suppose J is a member of a Banach algebra $\mathfrak{E}$ such that $\sigma(J)$ is an infinite set. The "Gelfand-transformation" $G(\mathfrak{E}, J)$ is the (unique) continuous homomorphism of $\mathfrak{E}[J]$ into $\boldsymbol{C}(\sigma(J))$ which maps J on I.*

2.13 *Remark.* If Φ is the Gelfand-transformation $G(\mathfrak{E}, J)$, then (by 2.11 and 2.10) Φ is a spectral homomorphism of $\mathfrak{E}[J]$ into $\boldsymbol{C}(\sigma(J))$ such that, if $\psi = [J;\mathfrak{E}]$, then

$$\Phi(\psi(P)) = P \qquad \text{for all } P \text{ in } \mathfrak{P}.$$

In other words, $G(\mathfrak{E}, J)$ is the continuous extension to $\mathfrak{E}[J]$ of the inverse of $[J;\mathfrak{E}]$.

3. The Banach algebras (*AC*) and (*BV*).

The set (BV) of functions of bounded variation on $[-\pi, \pi]$ becomes a Banach algebra under the norm

$$(4) \qquad \| B \|_b = \| B \|_\infty + \operatorname{var} B \qquad (B \varepsilon (BV))$$

(cf. [8], p. 448), where $\operatorname{var} B$ designates the total variation of B on $[-\pi, \pi]$ and

$$(5) \qquad \| B \|_\infty = \sup \{ | B(\theta) | : | \theta | \leqq \pi \}.$$

Banach and Mazur ([2], p. 101) make (BV) into a Banach algebra by adopting the norm $\| B \|_v$ obtained from (4) by replacing $\| B \|_\infty$ by $B(-\pi)$; note that $\| B \|_v \leqq \| B \|_b \leqq 2 \| B \|_v$, so that the resulting topologies are equivalent. The set (AC) of absolutely continuous members of (BV) forms a Banach algebra under the norm of (BV) (cf. [1], p. 194). It will be implied henceforth that (4) defines the norm for both (BV) and (AC). That (AC) forms then a separable space, follows from 3.1.

3.1 LEMMA. *If $A \varepsilon \mathfrak{J} = (AC)$, then there exists a sequence $\{P_n\}$ of members of $\mathfrak{P}$ such that $A = \mathfrak{J} \operatorname{lm} P_n$.*

Proof. Let P_n be the n-th Bernstein polynomial of A. Since A is continous on $[-\pi, \pi]$, we see from ([9], p. 5) that $\{P_n\}$ converges uniformly to A; therefore $\lim \| A - P_n \|_\infty = 0$. But $A \varepsilon (AC)$ and consequently (cf. [10], Satz 7) $\lim \operatorname{var}(A - P_n) = 0$. This concludes the proof, since (4) defines the norm of $\mathfrak{J}$.

4. The main results. Choose for p a fixed value $p > 1$, and denote by Ψ the mapping $F \to \langle \Lambda F \rangle_p$ defined in the introduction. In a recent paper [13], Stečkin has shown that Ψ maps (BV) into the Banach algebra $\mathfrak{E}$ of bounded operators on l_p. From now on, $\mathfrak{R}$ will denote either one af the two Banach algebras (BV) and (AC); thus, $A = \mathfrak{R} \operatorname{lm} P_n$ means that $\lim \| A - P_n \|_b = 0$. Henceforth, $\mathcal{C} = C([-\pi, \pi])$; thus, $A = \mathcal{C} \operatorname{lm} P_n$ means that $\lim \| A - P_n \|_\infty = 0$ (see § 2 and (5)).

4.1 LEMMA. *If $F_n \varepsilon \mathfrak{R}$ and $A = \mathfrak{R} \operatorname{lm} F_n$, then $A = \mathcal{C} \operatorname{lm} F_n$. Suppose moreover that $T = \mathfrak{E} \operatorname{lm} \Psi(F_n)$; then $T = \Psi(A)$.*

Proof. Set $f_n = A - F_n$ and note that $\| f_n \|_\infty \leqq \| f_n \|_b$; the conclusion $A = \mathcal{C} \operatorname{lm} F_n$ follows. The completeness of the space $\mathfrak{R}$ necessitates that $A \varepsilon \mathfrak{R} \subset (BV)$; the conclusion $T = \Psi(A)$ is now given by 4.3 in [6].

4.2 THEOREM. *The mapping Ψ is a continuous isomorphism of $\mathfrak{R}$ into $\mathfrak{E}$ such that $\Psi(I) = H$, where H is characterized by the Laurent matrix $(a_{n-\nu})$ with $a_m = i(-1)^m/m$ and $a_0 = 0$.*

Proof. It was shown in [6] that $\Psi(I) = H$ (the notation used there is $I_\#$ instead of H), and that Ψ is an isomorphism of the Banach space $\mathfrak{R}$ into the Banach space $\mathfrak{E}$. It will therefore suffice to show that Ψ is a closed operator (see [5], p. 30). To that effect, suppose $F_n \varepsilon \mathfrak{R}$, $A = \mathfrak{R} \operatorname{lm} F_n$, and $T = \mathfrak{E} \operatorname{lm} \Psi(F_n)$; the conclusion $T = \Psi(A)$ is given by 4.1.

4.3 THEOREM. *The isomorphism Ψ is the only continuous homomor-*

phism of $\mathfrak{J} = (AC)$ *into* $\mathfrak{E}$ *such that* I *is mapped on* H. *The image* $\Psi(\mathfrak{J})$ *is dense in* $\mathfrak{E}[H]$, *and the Gelfand-transformation* $G(\mathfrak{E}, H)$ *coincides on* $\Psi(\mathfrak{J})$ *with the inverse mapping of* Ψ. *Moreover, the properties* (i)-(iii) *of the introduction are satisfied.*

Proof. Suppose $\iota = 1, 2$ and let Ψ_ι be continuous homomorphisms of $\mathfrak{J}$ into $\mathfrak{E}$ such that $\Psi_\iota(I) = H$. The restrictions ψ_ι of Ψ_ι to $\mathfrak{P}$ are homomorphisms of $\mathfrak{P}$ into $\mathfrak{E}$ such that $\psi_\iota(I) = H$; having proved in ([6], 6.4) that $\sigma(H) = [-\pi, \pi]$, we can infer from 2.9 that

$$\psi_\iota = [H; \mathfrak{E}] = \psi \quad \text{and} \quad \Psi_\iota(P) = \psi(P) \qquad (\text{when } P \,\varepsilon\, \mathfrak{P}). \tag{6}$$

If $A \,\varepsilon\, \mathfrak{J}$, then (by 3.1) there exists a sequence $\{P_n\}$ satisfying

$$A = \mathfrak{J} \operatorname{lm} P_n \qquad (n \to \infty, P_n \,\varepsilon\, \mathfrak{P}). \tag{7}$$

From the continuity of Ψ_ι follows that $\Psi_\iota(A) = \mathfrak{E} \operatorname{lm} \Psi_\iota(P_n)$. This enables us to derive from (6) that

$$\Psi_\iota(A) = \mathfrak{E} \operatorname{lm} \psi(P_n) \qquad (n \to \infty). \tag{8}$$

Thus $\Psi_1(A) = \Psi_2(A)$, and $\Psi_1 = \Psi_2$. Hence, there is at most one continuous homomorphism of $\mathfrak{J}$ into $\mathfrak{E}$ which maps I on H; from 4.2 now follows that Ψ is the only such homomorphism. In the following, (6) and (8) should be viewed in the light of $\Psi_\iota = \Psi$.

Next, observe that $\psi(P_n)$ is in the closure $\mathfrak{E}[H]$ of $\psi(\mathfrak{P})$ (by 2.8), so that $\Psi(A) \,\varepsilon\, \mathfrak{E}[H]$ (from (8)). This implies $\Psi(\mathfrak{J}) \subset \mathfrak{E}[H]$. Since $\psi(\mathfrak{P}) \subset \Psi(\mathfrak{J})$ (by (6) and $\mathfrak{P} \subset \mathfrak{J}$), the denseness of $\psi(\mathfrak{P})$ in $\mathfrak{E}[H]$ now necessitates that

$$\Psi(\mathfrak{J}) \text{ is dense in } \mathfrak{E}[H].$$

We now turn to the conclusions involving $G(\mathfrak{E}, H)$. Again referring to [6] for the result $\sigma(H) = [-\pi, \pi]$, we see from 2.12 that the Gelfand-transformation $G(\mathfrak{E}, H)$ is a continuous homomorphism Φ of $\mathfrak{E}[H]$ into $\mathfrak{C} = \boldsymbol{C}([-\pi, \pi])$. A successive application of (8) with the continuity of Φ, and 2.13 (with $\psi = [H; \mathfrak{E}]$) yields

$$\Phi(\Psi(A)) = \mathfrak{C} \operatorname{lm} \Phi(\psi(P_n)) = \mathfrak{C} \operatorname{lm} P_n \qquad (n \to \infty).$$

This result, combined with the consequence $A = \mathfrak{C} \operatorname{lm} P_n$ of (7) (see 4.1), shows that $\Phi(\Psi(A)) = A$. Accordingly, if $T = \Psi(A)$ and if ϕ is the inverse of the mapping Ψ, then $\Phi(T) = A = \phi(T)$, which states that Φ coincides with ϕ on $\Psi(\mathfrak{J})$. On the other hand, $\Phi(T) = A \,\varepsilon\, \mathfrak{J}$ and 2.4 shows that (iv)-(v) are satisfied, in consequence of Φ being a spectral homomorphism

of $\mathfrak{X} = \mathfrak{C}[H]$ into $C(\mathfrak{S}), \mathfrak{S} = [-\pi, \pi]$ (see 2.13). Therefore (i)-(iii) hold, and the proof is completed.

4.4 *Remark.* Suppose $A \varepsilon (AC)$ and set $T = \Psi(A)$. In the preceding paragraph, we have pointed out that the relations 2.4 (iv) are satisfied; hence $\| A \|_\infty \leqq \| T \|$. In case $p = 2$, then $\| A \|_\infty = \| T \|$. This was proved by F. Riesz [12] and can in the present context be derived as follows. The operator T is a member of the abelian Hilbert space $\mathfrak{C}[H]$ and the same holds for its adjoint T^*; therefore T is a normal operator. But then $\| T \| = \sup \{ | \lambda | : \lambda \varepsilon \sigma(T) \} = \sup \{ | \lambda | : \lambda \varepsilon A(\mathfrak{S}) \} = \| A \|_\infty$; the first equality holds for any normal operator, the second follows from $\sigma(T) = A(\mathfrak{S})$, the third equality is obvious.

4.5 *Remark.* We here supply a few details connected with 2.4 and 2.6. It should be kept in mind that l_1 is a Banach algebra $\mathfrak{C}$ having a member J such that $\sigma(J) = \Gamma_1 = \{ \lambda : | \lambda | = 1 \}$, and l_1 is the subalgebra $\mathfrak{C}[J]$ generated in $\mathfrak{C}$ by J. The inverse ∇ of the mapping $f \rightarrow \Delta f$ is an isomorphism of l_1 onto $\mathcal{U}$ (the symbols Δf and $\mathcal{U}$ are defined in 1.1). If we set $\mathfrak{X} = l_1$, $\mathfrak{F} = \mathcal{U}$ and $\mathfrak{S} = \Gamma_1$, then (iv)-(v) hold when $T \varepsilon \mathfrak{X}$ and $A = \nabla(T)$. This can be seen as follows. Let $\Phi(T)$ be the function defined on Γ_1 by $f(\lambda) = \sum T_n \lambda^n$; clearly Φ is a continuous homomorphism of $l_1 = \mathfrak{C}[J]$ into $C(\mathfrak{S})$, and 2.12 now shows that Φ is the Gelfand-transformation $G(\mathfrak{C}, J)$. Accordingly, Φ is spectral; the conclusion is now obtained from 2.4 and the fact that $\nabla(T)$ coincides with $\Phi(T)$ on Γ_1.

Consequently, $f^{-1} \varepsilon \mathcal{U}$ when $f \varepsilon \mathcal{U}$ and provided $0 \notin f(\Gamma_1)$. This conclusion can be used in the otherwise obvious derivation of (i′)-(iii′) from (i)-(iii). Moreover, we have just seen that $\sigma(T) = (\Phi(T))(\Gamma_1)$; since the isomorphism V of 2.6 satisfies $(V(T))(\theta) = (\Phi(T))(e^{i\theta})$, we can conclude that V is a spectral isomorphism of $\mathfrak{X} = l_1$ onto the ring of all absolutely convergent Fourier series. In view of these facts, the classical Wiener theorems ([7], pp. 72-73) now appear as consequences of 2.4.

PURDUE UNIVERSITY.

REFERENCES.

[1] C. R. Adams and A. P. Morse, "On the space (BV)," *Transactions of the American Mathematical Society*, vol. 42 (1937), pp. 194-205.

[2] S. Banach and S. Mazur, "Zur Theorie der linearen Dimension," *Studia Mathematica*, vol. 4 (1933), pp. 100-112

[3] N. Dunford, "Spectral theory I. Convergence to projections," *Transactions of the American Mathematical Society*, vol. 54 (1943), pp. 185-217.

[4] P. Hartman and A. Wintner, "On the spectra of Toeplitz's matrices," *American Journal of Mathematics*, vol. 72 (1950), pp. 359-366.

[5] E. Hille, "Functional analysis and semi-groups," *American Mathematical Society Colloquium Publications*, vol. 31, New York (1948).

[6] G. L. Krabbe, "Abelian rings and spectra of operators on l_p," *Proceedings of the American Mathematical Society.*

[7] L. H. Loomis, *An introduction to abstract harmonic analysis*, D. Van Nostrand Company, New York (1953).

[8] E. R. Lorch, "The structure of normed abelian rings," *Bulletin of the American Mathematical Society*, vol. 50 (1944), pp. 447-463.

[9] G. G. Lorentz, *Bernstein Polynomials*, University of Toronto Press, Toronto (1953).

[10] ———, "Zur Theorie der Polynome von S. Bernstein," *Matematicheskii Sbornik*, vol. 44 (1937), pp. 543-556.

[11] J. D. Newburgh, "The variation of spectra," *Duke Mathematical Journal*, vol. 18 (1951), pp. 165-176, in particular, p. 169.

[12] F. Riesz, *Les systèmes d'équations linéaires à une infinité d'inconnues*, Paris (1913), pp. 154-155 and pp. 171-175.

[13] S. B. Stečkin, "On bilinear forms," *Comptes Rendus (Doklady) de l'Academie des Sciences de l'URSS*, N. S. 71 (1950), pp. 237-240.

[14] O. Toeplitz, "Zur Theorie der quadratischen Formen von unendlichvielen Veränderlichen," *Nachrichten von der Gesellschaft der Wissenschaften zu Göttingen* (1910), pp. 489-506.

[15] ———, "Zur Theorie der quadratischen und bilinearen Formen von unendlichvielen Veränderlichen," *Mathematische Annalen*, vol. 70 (1911), pp. 351-376, in particular, pp. 354-360.

[16] ———, "Zur Transformation der Scharen bilinearer Formen von unendlichvielen Veränderlichen," *Nachrichten von der Gesellschaft der Wissenschaften zu Göttingen* (1907), pp. 110-115.

A NOTE ON THE CLASS-NUMBERS OF ALGEBRAIC NUMBER FIELDS.*

By N. C. Ankeny, R. Brauer and S. Chowla.

Let F be an algebraic number field of finite degree n over the field P of rational numbers. Denote by $h(F)$ and $d(F)$ the class-number and discriminant of F respectively. According to Minkowski each class of ideals of F contains an ideal of normal at most $|d(F)|^{\frac{1}{2}}$. Landau [1] used this result to deduce

$$h(F) < c_1 \, |d(F)|^{\frac{1}{2}} (\log |d(F)|)^{n-1}$$

where c_1 is a constant depending on n alone. If ϵ is a given positive number, this implies

$$h(F) < c_2 \, |d(F)|^{\frac{1}{2}+\epsilon} \tag{1}$$

where the constant c_2 depends on n and ϵ.

We shall show that, for suitable fields F, the rather rough estimate (1) is actually remarkably sharp. Indeed, given any positive integer $n \geqq 2$, let r_1 and r_2 be any two non-negative integers such that $r_1 + 2r_2 = n$. We shall prove that for every $\epsilon > 0$ there exist infinitely many fields F which have exactly r_1 real and $2r_2$ imaginary conjugate fields and are such that

$$h(F) > |d(F)|^{\frac{1}{2}-\epsilon} \tag{2}$$

holds.

For the proof, we use the following result of R. Brauer [2] which confirmed a conjecture of C. L. Siegel. For all fields F of given degree n with sufficiently large $|d(F)|$, we have

$$h(F)R(F) > |d(F)|^{\frac{1}{2}-\delta} \tag{3}$$

where $R(F)$ is the regulator of F and $\delta > 0$ a given arbitrary constant. For $n = 2$, this had already been proved by Siegel.

We can immediately settle the case $n = 2$ as follows. If F is an imaginary quadratic extension of the field P of rationals, then (2) holds

* Received May 24, 1955; revised July 18, 1955.

for all fields F with $|d(F)|$ exceeding a certain limit depending on ϵ alone. To construct real quadratic fields with the property in question, let F be generated by $[m^2+1]^{\frac{1}{2}}$, where we select the rational integer $m \geqq 2$ so that m^2+1 is square free. We recall the well-known result of Estermann [4] that m^2+1 is square-free for infinitely many choices of m. Since $E = m + [m^2+1]^{\frac{1}{2}}$ is a unit contained in F and since $d(F) = m^2+1$ or $d(F) = 4(m^2+1)$, we have

$$R(F) \leqq \log(m + [m^2+1]^{\frac{1}{2}}) \leqq \log(2[m^2+1]^{\frac{1}{2}}) \leqq \log(m^2+1) \leqq \log d(F).$$

So (2) follows from (3) for the fields $F = P([m^2+1]^{\frac{1}{2}})$, where m^2+1 is square-free and sufficiently large.

Section 1. n is any fixed integer $\geqq 3$, N is an arbitrary positive integer that is taken to be sufficiently large.

Let $a_1, a_2, a_3, \cdots, a_{n-1}$ be arbitrary fixed distinct integers, and take $a_n = N$. We define

$$f_N(x) = \prod_{j=1}^{n} (x - a_j) + 1.$$

We prove (see the references [5], [6]).

LEMMA 1. *If N is sufficiently large, $f_N(x)$ enjoys the following properties:*

(4) $\qquad f_N(a_j) = 1 \qquad (1 \leqq j \leqq n).$

(5) $\qquad$ *$f_N(x) = 0$ has n distinct real roots: $\theta_N^{(1)}, \theta_N^{(2)}, \cdots, \theta_N^{(n)}$.*

For a suitable arrangement of these roots there exist constants $b_i \neq 0$ such that, for $N \to \infty$,

(6) $\quad N(\theta_N^{(i)} - a_i) \to b_i \neq 0 \quad (1 \leqq i \leqq n-1), \qquad \theta_N^{(n)} = N + O(N^{-1}).$

(7) *$f_N(x)$ is an irreducible equation of degree n in the field of rational numbers.*

Proof. To prove that the roots are real, we note that $f(a_j + \frac{1}{2})$ and $f(a_j - \frac{1}{2})$, $j = 1, \cdots, n$ are of different sign for N sufficiently large.

(4) is clear from the definition of $f_N(x)$. Next, *as* $N \to +\infty$,

$$-(1/N)f_N(x) \to \prod_{i=1}^{n-1} (x - a_i).$$

Because of the continuity of the roots as functions of the parameter N, each of the a_i, $(i \neq n)$, is the limit of a root $\theta_N^{(i)}$ of $f_N(x)$. Also

$$(\theta_N^{(i)} - a_i)(\theta_N^{(i)} - N)Z + 1 = 0$$

where, as $N \to \infty$,

$$Z = \prod_{\mu \neq i,\, \mu \neq n} (\theta_N^{(i)} - a_\mu) \to \prod_{\mu \neq i,\, n} (a_i - a_\mu) \neq 0.$$

Hence $N(\theta_N^{(i)} - a_i) \to \prod_{\mu \neq i,\, n} (a_i - a_\mu)^{-1}$, and this limit is a constant $b_i \neq 0$. In particular, $\theta_N^{(i)} = a_i + O(N^{-1})$ $(1 \leqq i \leqq n-1)$. Since $\sum_{i=1}^{n} \theta_N^{(i)} = \sum_{i=1}^{n} a_i$, it follows that $\theta_N^{(n)} = a_n + O(N^{-1}) = N + O(N^{-1})$. The above equations prove (6).

Next let us assume that there are infinitely many N for which $f_N(x)$ is reducible and set $f_N(x) = g_{1,N}(x) g_{2,N}(x)$ where $g_{1,N}(x)$ and $g_{2,N}(x)$ have integral rational coefficients. By (5), we may always select $g_{1,N}(x)$ such that $\theta_N^{(n)}$ is not a root of $g_{1,N}(x)$. If $\theta_N^{(i)}$ is a root of $g_{1,N}(x)$, all the conjugates of $N - \theta_N^{(i)} = a_n - \theta_N^{(i)}$ will be larger than 1 for sufficiently large N. This is a contradiction, since (4) shows that $a_n - \theta_N^{(i)}$ is a unit and hence has norm ± 1. This completes the proof of the lemma.

We note that the coefficients of $f_N(x)$ are linear functions of N. Hence $\operatorname{discr}(f_N(x)) = g(N)$, where $g(N)$ is a polynomial in N whose coefficients are rational integers. On considering the order of magnitude of

$$g(N) = \prod_{i>j} (\theta_N^{(i)} - \theta_N^{(j)})^2,$$

we see that $g(N)$ is of degree $2(n-1)$.

The following lemma is proved for general polynomials $g(x)$ and later specialized to the $g(x)$ above. We prove

LEMMA 2. *Let $g(x)$ be a polynomial with integral coefficients of degree $s > 0$. Let m be the greatest common divisor of the values of $g(x)$ for integral x. Let $\rho \geqq 3$ be a fixed number, and let U be a number chosen sufficiently large. If U^* denotes the number of integers N for which $U < N \leqq 2U$ and $g(N)/m$ has all prime factors greater than $V = (\frac{1}{4} \log U)^\rho$, then*

(8) $$U^* > cU/(\log \log U)^s.$$

Here c is an absolute positive constant depending only on the coefficients of $g(x)$ and on ρ.

The proof of Lemma 2 is in turn based on the following Lemma 3 which is a recent theorem of de Bruijn [3]. For a simple proof, see the paper of W. E. Briggs and S. Chowla [7].

LEMMA 3. *Let $f(x,y)$ denote the number of positive integers $\leqq x$ all of whose prime factors are $\leqq y$. Then*

$$f(x, (\log x)^{\rho}) = O(x^{1-1/(2\rho)}) \tag{9}$$

where $\rho > 2$ and the constant implied in the O-symbol depends on ρ alone.

Section 2. In this section, we shall prove Lemma 2.

For a positive integer d, let $\lambda(d)$ denote the number of solutions of $g(x) \equiv 0 \pmod d$ with $0 \leqq x < d$. In the following, the letter p will always denote a prime number. We have

$$\lambda(p) \leqq s \text{ if } p \nmid m, \qquad \lambda(d) \leqq d \text{ for all } d. \tag{10}$$

Denote by $\nu(d)$ the number of distinct prime factors of d. Then, for any fixed $\epsilon > 0$,

$$s^{\nu(d)} = O(d^{\epsilon}). \tag{11}$$

To prove this, take $s > 1$ and note that

$$s^{\nu(d)} = 2^{\nu(d)\log s/\log 2} \leqq \{\tau(d)\}^{\log s/\log 2}$$

where $\tau(d)$ is the number of divisors of d. Since (see, for example, Landau's Vorlesungen ü. Zahlentheorie)

$$\tau(d) = O(d^{\theta}), \qquad \theta = \epsilon \log 2/\log s;$$

(11) now follows. We shall use the symbol $\mu(t)$ for the well-known Möbius function so that

$$\mu(t) = 0 \text{ if } p^2 \mid t \text{ for some prime } p, \qquad \mu(t) = (-1)^{\nu(t)} \text{ for square-free } t.$$

As is well-known,

$$\lambda(ab) = \lambda(a)\lambda(b) \text{ for } (a,b) = 1. \tag{12}$$

We also need

$$\lambda(d) = O(d^{\epsilon}) \qquad (d \to \infty), \tag{13}$$

if d is a multiple of m such that d/m is square-free. This can be seen easily (actually (13) holds for all d).

We now introduce the function $Q(d)$ defined as follows:

(i) $Q(d) = \mu(d)$ if $1 \leqq d \leqq U^{\frac{1}{4}}$.

(ii) $Q(d) = \mu(d)$ if $U^{\frac{1}{4}} < d \leqq U^{\frac{1}{2}}$ and $\nu(d)$ odd.

(iii) $Q(d) = 0$ if $U^{\frac{1}{4}} < d \leqq U^{\frac{1}{2}}$ and $\nu(d)$ even.

(iv) $Q(d) = 0$ if $d > U^{\frac{1}{2}}$.

Observe that $Q(d) = 0$, if $\mu(d) = 0$ and that $Q(d) = 1$, if and only if $d \leqq U^{\frac{1}{4}}$, d square-free, $\nu(d)$ even. Further, set $Q(d) = 0$, if d is not integral.

Lemma 4. *If $k > 0$ is an integer and if $S(k) = \sum_{d \mid k} Q(d/m)$, then*

(α) *$S(k) = 1$ if $k/m = k_0$ is integral and not divisible by a prime $\leqq U^{\frac{1}{2}}$;*

(β) *$S(k) \leqq 0$ in all other cases.*

Proof. If $k \not\equiv 0 \pmod{m}$, obviously $S(k) = 0$. Suppose that $k_0 = k/m$ is integral. Then

$$S(k) = \sum_{t \mid k_0} Q(t). \tag{14}$$

If k is not divisible by primes $\leqq U^{\frac{1}{2}}$, then $S(k) = Q(1) = 1$. Suppose that k_0 is divisible by a prime $p \leqq U^{\frac{1}{2}}$ and choose p as the least such prime. Since it suffices to take square-free t in (14), we can arrange these t in pairs h and hp where h is a divisor of k_0 which is prime to p. In order to prove (β) it will be sufficient to show that $Q(h) + Q(hp) \leqq 0$. If this was not so, we must have either

$$Q(hp) = 1, \quad Q(h) \geqq 0 \quad \text{or} \quad Q(h) = 1, \quad Q(hp) = 0.$$

Since $Q(x) = 1$ implies that $x \leqq U^{\frac{1}{4}}$ and that x is square-free, it follows easily from $Q(hp) = 1$ that $Q(h) = -1$ and hence the former case is impossible. Suppose then that $Q(h) = 1$, $Q(hp) = 0$. Then $h \leqq U^{\frac{1}{4}}$, h square-free and $\nu(h)$ even. In order to have $Q(hp) = 0$, we must have $hp > U^{\frac{1}{2}}$ and hence $p > U^{\frac{1}{4}}$. Then every divisor $h \neq 1$ of k_0 with $\nu(h)$ even would exceed $U^{\frac{1}{2}}$ and this would imply $Q(h) = 0$. Hence $h = 1$. In this case $Q(hp) = Q(p) = -1$. This is a contradiction and the lemma is proved.

Suppose that U is so large that $V = (\frac{1}{4} \log U)^{\rho} \leqq U^{\frac{1}{2}}$ and that $V \geqq 2m$. Let U^* denote the number of N with $U < N \leqq 2U$ for which $k_1 = g(N)/m$ is not divisible by primes $\leqq V$. Then if $(g(N), [V]!)$ denotes the g.c.d. of $g(N)$ and $[V]!$.

$$U^* \geqq \sum_{U < N \leqq 2U} S((g(N), [V]!)). \tag{15}$$

Indeed, by Lemma 4 all terms here are either $\leqq 0$ or have the value 1, and the latter case arises only if $k = (g(N), [V]!) = mk_0$ where k_0 has no prime factors $\leqq U^{\frac{1}{2}}$. In this case, $g(N)/m$ cannot have a prime factor $\leqq V$ and all such N are certainly counted by U^*.

Let g_d denote the largest prime dividing any integer $d > 1$ and set $g_1 = 1$. Then

$$U^* \geqq \sum_{U<N\leqq 2U} \sum_{d\,|\,(g(N),[V]!)} Q(d/m) = \sum_d Q(d/m) \sum_N 1.$$

Here, d ranges over all positive integers such that

(a) $g_d \leqq V$; (b) $d \equiv 0 \pmod{m}$; (c) d/m is square-free;

while N ranges over all integers with $U < N \leqq 2U$ and $g(N) \equiv 0 \pmod{d}$. Thus,

$$U^* \geqq U \sum_d Q(d/m)\lambda(d)/d - \sum_d |Q(d/m)|\lambda(d).$$

Let $\epsilon < 1/(2\rho)$ be a positive constant. Using (13) and recalling that $Q(d) = 0$ for $d > U^{\frac{1}{2}}$, we obtain

$$U^* \geqq U \sum_d Q(d/m)\lambda(d)/d - O(U^{\frac{1}{2}+\epsilon}). \tag{16}$$

If we replace $Q(d/m)$ by $\mu(d/m)$, the error term is

$$\left|\sum_d Q(d/m)\lambda(d)/d - \mu(d/m)\lambda(d)/d\right| \leqq \sum_t \lambda(t)/t \leqq O(\sum_t t^{-1+\epsilon}) \tag{17}$$

where t ranges over all integers with $t > mU^{\frac{1}{4}}$ for which $g_t \leqq V$. The convergence of the sum on the right will become evident from the following argument.

For $x > mU^{\frac{1}{4}} \geqq U^{\frac{1}{4}}$, we have $V = (\frac{1}{4}\log U)^\rho < (\log x)^\rho$. If $f(x,y)$ is the expression introduced in Lemma 3, this implies $f(x,V) \leqq f(x,(\log x)^\rho) = O(x^{1-1/(2\rho)})$. Now, if n ranges over the integers larger than $mU^{\frac{1}{4}}$, we find by partial summation

$$\sum_t t^{-1+\epsilon} = \sum_n n^{-1+\epsilon}(f(n,V) - f(n-1,V))$$

$$\leqq (1-\epsilon)\int_{mU^{1/4}}^{\infty} f(x,V)x^{-2+\epsilon}dx = O\left(\int_{mU^{1/4}}^{\infty} x^{-1-1/(2\rho)+\epsilon}dx\right),$$

$$\sum_t t^{-1+\epsilon} = O(U^{-1/(8\rho)+\epsilon/4}). \tag{18}$$

Since $\epsilon < 1/(2\rho)$, the exponent of U is negative. Hence, by substituting (17) and (18) in (16), we have

$$U^* \geqq U \sum_d \mu(d/m)\lambda(d)/d - O(U^{1-\gamma}) \tag{19}$$

where γ is a positive constant and where d still ranges over the integers which satisfy the conditions (a), (b), and (c).

Set $d = m\alpha\beta$ with $(\alpha, m) = 1, \beta \mid m$. Then

$$\mu(d/m)\lambda(d)/d = \mu(\alpha)\lambda(\alpha)/\alpha\, \mu(\beta)\lambda(\beta m)/\beta m)$$

and hence

$$\sum_d \mu(d/m)\lambda(d)/d = \sum_\alpha \mu(\alpha)(\lambda(\alpha)/\alpha) \sum_{\beta \mid m} \mu(\beta)\lambda(\beta m)/(\beta m). \tag{20}$$

In the first sum on the right, α ranges over the integers prime to m for which $g_\alpha \leqq V$. Therefore

$$\sum_\alpha \mu(\alpha)\lambda(\alpha)/\alpha = \prod_{p \leqq V,\, p \nmid m} (1 - \lambda(p)/p).$$

If β is a fixed divisor of m, let $m_1(\beta)$ denote the product of those prime powers of m which are prime to β and set $m = m_1(\beta)m_2(\beta)$. Then $\lambda(\beta m) = \lambda(m_1(\beta))\lambda(\beta m_2(\beta))$. Since $m_1(\beta) \mid m$, we have $\lambda(m_1(\beta)) = m_1(\beta)$. Thus,

$$\sum_\beta \mu(\beta)\lambda(\beta m)/(\beta m) = \sum_\beta \mu(\beta)\lambda(\beta m_2(\beta))/(\beta m_2(\beta)).$$

If $\beta = \beta'\beta''$ with $(\beta', \beta'') = 1$, then $m_2(\beta) = m_2(\beta')m_2(\beta'')$ and

$$\lambda(\beta m_2(\beta)) = \lambda(\beta' m_2(\beta'))\lambda(\beta'' m_2(\beta'')).$$

Hence

$$\sum_\beta \mu(\beta)\lambda(\beta m)/(\beta m) = \prod_{p \mid m} (1 - \lambda(p m_2(p))/(p m_2(p))).$$

Now, $pm_2(p)$ does not divide m and, consequently, $\lambda(pm_2(p)) \neq pm_2(p)$. It follows that the last product is a positive constant c_3. This shows that

$$\begin{aligned}\sum_d \mu(d/m)\lambda(d)/d &= c_3 \prod_{p \leqq V,\, p \nmid m} (1 - \lambda(p)/p) \\ &\geqq c_3 \prod_{p \leqq s} (1 - (p-1)/p) \prod_{s < p \leqq V} (1 - s/p).\end{aligned}$$

By Mertens' theorem on prime numbers, we have

$$\sum_d \mu(d/m)\lambda(d)/d \geqq c_4(\log V)^{-s} \geqq c_5(\log\log U)^{-s} \tag{21}$$

where c_4 and c_5 are also positive constants, independent of U. On combining (19) and (21), we obtain $U^* \geqq cU(\log\log U)^{-s}$ with a positive constant c, and this proves Lemma 2.

Section 3. We shall now prove that there are infinitely many totally real algebraic number fields of degree $n \geqq 3$ over the rationals, and such that (2) holds.

Writing $F_N = P(\theta_N^{(1)})$ we prove

LEMMA 5. *There exist an infinite set of N for which* $|d(F_N)| > (\frac{1}{4}\log\frac{1}{2}N)^{\rho}$ *where* ρ *is an arbitrary but fixed positive integer.*

We refer back to Lemma 2. Let $S(U)$ denote the set of N which satisfy the hypotheses of Lemma 2. Now we observe that $d(F_N)$ divides the discriminant $g(N)$ of $f_N(x)$. For $N \varepsilon S(U)$ it follows that if $|d(F_N)| \leqq (\frac{1}{4}\log\frac{1}{2}N)^{\rho}$, then $|d(F_N)| \leqq (\frac{1}{4}\log U)^{\rho}$ and $d(F_N)\,|\,m$. Thus if Lemma 5 were not true, we would have $d(F_N)\,|\,m$ for all $N \varepsilon S(U)$, U sufficiently large. It is a well-known theorem of Minkowski that there are only a finite number of algebraic number fields whose discriminant has a presecribed value. Hence there are only a finite number of fields whose discriminant divides m.

By Lemma 2 as U increases there is an increasing number of $N \varepsilon S(U)$. Let us select U sufficiently large so that at least one of the fields, say F', contains $\theta_N^{(1)}$ for at least $(n-1)!+1$ values of N belonging to $S(U)$. Consider the n isomorphisms of F' onto its conjugate fields. Using Dirichlet's chest of drawers principle we shall be able to select N_1 and N_2 with $U < N_1 < N_2 \leqq 2U$ such that $\theta_{N_1}^{(1)}$, $\theta_{N_2}^{(1)} \varepsilon F'$ and that each of the n isomorphisms of F' maps $\theta_{N_1}^{(1)}$, $\theta_{N_2}^{(1)} \varepsilon F'$ on $\theta_{N_1}^{(j)}$, $\theta_{N_2}^{(j)}$ respectively, with the same j. Write $\phi^{(j)} = \theta_{N_1}^{(j)} - \theta_{N_2}^{(j)}$. Now $\phi^{(1)} \varepsilon F'$ and the $\phi^{(j)}$ are the n conjugates of $\phi^{(1)}$. For $N_1 \neq N_2$ we have $\theta_{N_1}^{(1)} \neq \theta_{N_2}^{(1)}$ since we have the roots of two different irreducible equations. Hence $\phi^{(1)} \neq 0$. Now by Lemma 1, $|\phi^{(j)}| < c_6/N_1 \leqq c_6/U$ for $j = 1, 2, 3, \cdots, n-1$.[1] Further $|\phi^{(n)}| < N_2 + c_7 \leqq 2U + c_7$. Hence

$$|N_{F',P}(\phi^{(1)})| = |\prod_{j=1}^{n} \phi^{(j)}| < c_8 U^{-n+2}.$$

If U is sufficiently large, this implies $N_{F',P}(\phi^{(1)}) = 0$, since the norm of $\phi^{(1)}$ is an integer. Hence $\phi^{(1)} = 0$, a contradiction. This concludes the proof of Lemma 5.

We now prove

THEOREM 1. *For given* $\epsilon > 0$, *there exists an infinity of totally real fields* F *of given degree* n *such that the class-number* $h(F)$ *satisfies the inequality* $h(F) \geqq |d(F)|^{\frac{1}{2}-\epsilon}$, *where* $d(F)$ *is the discriminant of* F.

Proof. We show that we may take F as one of the infinitely many fields F_N where N satisfies Lemma 5 for a suitable choice of ρ and is sufficiently large. It is clear from the definition of $f_N(x)$ that

$$\theta_N^{(1)} - a_1, \theta_N^{(1)} - a_2, \cdots, \theta_N^{(1)} - a_{n-1}$$

[1] We denote by $c_6, c_7 \cdots$ positive constants which may depend on $a_1, a_2, \cdots, a_{n-1}$, but are independent of N and U.

are all units contained in F_N. The determinant

$$\Delta = \det(\log|\theta_N^{(j)} - a_i|), \qquad [1 \leqq j, i \leqq n-1],$$

is not 0 for sufficiently large N, as by (6) of Lemma 1 the non-diagonal terms are bounded, whereas the diagonal terms $\to -\infty$ for $N \to +\infty$. Hence $\theta_N^{(1)} - a_1, \theta_N^{(1)} - a_2, \cdots, \theta_N^{(1)} - a_{n-1}$ are multiplicatively independent in F_N. The same argument yields

$$|\Delta| \leqq c_9(\log N)^{n-1}. \tag{21}$$

Now F_N is a totally real field of degree n over the rationals. So the regulator $R(F)$ is at most equal to the regulator of a set of $n-1$ independent units. Hence, by (21) $R(F_N) \leqq c_9(\log N)^{n-1}$. Take $\delta = \epsilon/2$ and choose ρ in Lemma 5 so that $\rho\delta > n-1$. As N satisfies Lemma 5, $|d(F_N)| > (\frac{1}{4}\log\frac{1}{2}N)^\rho$, so, for sufficiently large N,

$$R(F_N) \leqq c_9(\log N)^{n-1} \leqq |d(F_N)|^\delta.$$

Hence from (3), it follows that $h(F_N) > |d(F_N)|^{\frac{1}{2}-2\delta}$ which completes the proof of Theorem 1.

Section 4. We now carry over the proof to the case when F is not totally real. Let $n > 2, r_1 \geqq 0, r_2 \geqq 0$ be given integers such that $n = r_1 + 2r_2$. We prove

THEOREM 1*. *There exist infinitely many algebraic number fields K of degree n over the field P of rational numbers such that K has r_1 real conjugates and $2r_2$ non-real conjugates with the following property. The class-number h lies above $|d|^{\frac{1}{2}-\epsilon}$, d the discriminant, $\epsilon > 0$ a given constant.*

1. Let the a_λ, $(1 \leqq \lambda \leqq r_1)$, and the $a_\mu > 0$, $(r_1 + 1 \leqq \mu \leqq r+1)$, be distinct integers, where, as usual, $r = r_1 + r_2 - 1$. It can now be assumed that $r_2 > 0$, i. e. $r \geqq r_1$. We take $a_{r+1} = N$. Then define

$$f(x) = \prod_{\lambda=1}^{r_1}(x - a_\lambda) \prod_{\mu=r_1+1}^{r+1}(x^2 + a_\mu) + 1. \tag{22}$$

so that, for $N \to +\infty$,

$$f(x)/N \to \prod_\lambda (x - a_\lambda) \prod_{\mu \neq r+1} (x^2 + a_\mu). \tag{23}$$

We can denote r roots of $f(x)$ by $\theta_N^{(j)}$, $1 \leqq j \leqq r$, in such a fashion that

$$\theta_N^{(j)} \to a_j \quad (j = 1, 2, \cdots, r_1), \qquad \theta_N^{(j)} \to ia_j^{\frac{1}{2}} \quad (j = r_1 + 1, \cdots, r).$$

Then there exist constants $b_j \neq 0$ such that

(24) $N(\theta_N^{(j)}-a_j)\to b_j,\ (1\leqq j\leqq r_1);\ N(\theta_N^{(j)2}+a_j)\to b_j,\ (r_1+1\leqq j\leqq r)$.

Actually,

$$b_j=-\prod_{\lambda\neq j}(a_j-a_\lambda)^{-1}\prod_{\mu\neq r+1}(a_j^2+a_\mu)^{-1} \qquad (j\leqq r_1);$$

$$b_j=-\prod_{\lambda}(ia_j^{\frac{1}{2}}-a_\lambda)^{-1}\prod_{\mu\neq j,r+1}(-a_j+a_\mu)^{-1}, \qquad (r_1+1\leqq j\leqq r).$$

For N sufficiently large, $\theta_N^{(j)}$, $1\leqq j\leqq r_1$, is real and $\theta_N^{(j)}$, $r_1+1\leqq j\leqq r_1+r_2-1$, has positive imaginary part. Then r_2-1 further roots are obtained in the form $\bar{\theta}_N^{(j)}$, $(r_1+1\leqq j\leqq r_1+r_2-1)$. These have the limits $-ia_j^{\frac{1}{2}}$. Since the limit in (23) has degree $n-2$, two of the roots of $f(x)$ tend to ∞. We see easily that these roots are not real for N sufficiently large and if $\theta_N^{(r+1)}$ is the one which has positive imaginary part, then $\theta_N^{(r+1)}=iN^{\frac{1}{2}}+O(1)$. (Actually one can show without difficulty that $\theta_N^{(r+1)}=iN^{\frac{1}{2}}+O(N^{-\frac{1}{2}(n-1)})$.

It follows from (22) that the r_1+r_2 elements

$$\theta-a_\lambda, \qquad \theta^2+a_\mu \tag{25}$$

are units of the field $K=P(\theta)$, θ one of the θ_j.

2. We show that $f(x)$ is irreducible in P for sufficiently large N. If this was not so, let $f_0(x)$ be an irreducible factor in P such that $\theta_N^{(r+1)}$ is not a root of $f_0(x)$. Then $\bar{\theta}_N^{(r+1)}$ is not a root either. Let θ be a root of $f_0(x)$. All the conjugates of θ are close to fixed values. Take the unit $(\theta^2+a_{r+1})^{-1}$. All its conjugates are less than 1 in absolute value for N sufficiently large. This is impossible. Thus $f(x)$ is irreducible in P. Hence $K=P(\theta)$ is a field of the given degree n with the given number r_1 of real conjugates.

3. Form the regulator R_0 of the r units (25) with $\mu\neq r+1$ of the field $F_N=P(\theta)$, using the conjugates corresponding to $\theta_N^{(1)},\cdots,\theta_N^{(r)}$. It follows from (24) that $|R_0|=O((\log N)^r)$. Hence $|R(F_N)|=O((\log N)^r)$. Now everything works as in the totally real case.

THE JOHNS HOPKINS UNIVERISTY, HARVARD UNIVERSITY, AND THE UNIVERSITY OF COLORADO.

REFERENCES.

[1] E. Landau, *Abschätzungen von Charaktersummen, Einheiten und Klassenzahlen*, Nachrichten von der Königlichen Gesellschaft der Wissenschaften zu Göttingen, 1918, pp. 79-97.

[2] R. Brauer, "On the zeta functions of algebraic number fields," *American Journal of Mathematics*, vol. 69 (1947), pp. 243-250.

[3] N. G. de Bruijn, "On the number of positive integers $\leqq X$ and free of prime factors $> Y$," *Koninklijke Nederlandse Akademie van Wetenschappen, Proceedings.* Series A, vol. 54 (1951), pp. 50-60.

[4] I. Estermann, "Einige Sätze über quadratfreie Zahlen," *Mathematische Annalen*, vol. 105 (1931), pp. 653-662.

[5] Pólya-Szegö, *Aufgaben und Lehrsätze aus der Analysis*, Berlin, J. Springer, 1925, p. 136.

[6] A. Brauer, R. Brauer, H. Hopf, "Über die Irreduzibilitat einige spezieller Klassen von Polynomen," *Jahnesbericht der Deutschen Mathematiker-Vereinigung*, vol. 35 (1926), pp. 99-112.

[7] S. Chowla and W. E. Briggs, "On the number of positive integers $\leqq x$ all of whose prime factors are $\leqq y$," *Proceedings of the American Mathematical Society*, vol. 6 (1955), pp. 558-562.

THE REPRESENTATION OF INTEGERS BY CERTAIN RATIONAL FORMS.*

By E. G. Straus and J. D. Swift.

1. Introduction. In a previous paper [2] we discussed the representation of integers by $f(x,y) = (ax^2 + bxy + cy^2)/(p + qxy)$ where $a \mid (b,q)$ $c \mid (b,q)$. Similar questions have been discussed by other authors, [1], [3], [4]. In this paper we intend to analyze the underlying ideas and to extend their applications.

We wish to investigate an algebraic Diophantine equation in $n+1$ unknowns which is of degree no higher than the second in every unknown and of first degree in at least one of the unknowns. We distinguish the latter unknown by calling it z, and denote the other unknowns by $x_1, \cdots, x_n$ writing $x = (x_1, \cdots, x_n)$. Solving for z we obtain

$$(1) \qquad f(x) = z, \qquad f(x) = N(x)/D(x),$$

where $N(x)$, $D(x)$ are polynomials in x of degree no higher than 2 in each x_i. We shall be concerned here with the case $\deg D \geqq \deg N$.

Our main results are *finiteness results*. More precisely, we shall see that in certain cases there is a finite number of infinite classes of solutions of (1) which correspond to solutions of simpler Diophantine equations obtained from (1) by replacing some of the x_i by functions of the other x_j's. These we shall call the *regular* solutions of (1). The solutions which are not contained in the regular classes are called *exceptional*. They will be finite in number if certain divisibility conditions are satisfied. To a more limited extent we shall also be able to obtain *infinity results*; that is, prove in some cases that there is an infinity of exceptional solutions, if the divisibility conditions are violated.

The method of attack is a combination of two extremely simple ideas described in Sections 2, 3, 4. The remaining sections are devoted to a more complete discussion of special cases for the purpose of illustration. Finally, we shall discuss some possible extensions of our method.

* Received October 25, 1954.

2. The critical cone. For the results in this section we need only the assumption $\deg D \geqq \deg N$ with no restriction on the degree in the individual x_i.

Definition. Let $D_1(x)$ be the homogeneous polynomial consisting of the terms of highest degree of $D(x)$. Then the *critical cone,* $\mathcal{C}$, of (1) is the locus

$$D_1(x) = 0. \tag{2}$$

A *conical neighborhood* of $\mathcal{C}$ is an open set of rays originating at the origin and containing $\mathcal{C}$.

LEMMA 1. *Let $\mathcal{K}$ be a conical neighborhood of $\mathcal{C}$, and let* $\deg D \geqq \deg N$. *Then* (1) *has at most a finite number of integral z for lattice points $x \notin \mathcal{K}$.*

Proof. Since $D_1(x) \neq 0$ in the exterior of $\mathcal{K}$ and since $\deg N \leqq \deg D$ there exist a radius r and a number M so that for $|x| = (x_1^2 + \cdots + x_n^2)^{\frac{1}{2}} > r$ and $x \notin \mathcal{K}$ we have $|f(x)| \leqq M$. On the other hand there is only a finite number of lattice points with $|x| \leqq r$.

COROLLARY. *If $D_1(x)$ is definite then $f(x)$ represents at most a finite number of integers.*

3. Conjugate points. For the results of this section we need only the assumption that N and D are at most quadratic in the x_i under consideration, with no restriction on their degrees.

Definition. If (1) is quadratic in x_i, then corresponding to each solution $(x_1, \cdots, x_i, \cdots, x_n, z)$ there is the *i-conjugate* solution $(x_1, \cdots, x'_i, \cdots, x_n, z)$, where x'_i is the conjugate of x_i when we consider (1) as an equation in x_i with all other unknowns fixed. We denote the i-conjugate of x by $x^{(i)}$.

A point x' is a conjugate of x if there exists a sequence $i_1, \cdots, i_k$ so that $x' = x^{(i_1)(i_2)\cdots(i_k)}$.

In general the conjugate of a lattice point has rational coordinates but is not a lattice point. However, writing

$$N(x) = a_i x_i^2 + b_i x_i + c_i, \qquad D(x) = A_i x_i^2 + B_i x_i + C_i,$$

where a_i, b_i, c_i, A_i, B_i, C_i are independent of x_i, we obtain from (1)

$$(a_i - zA_i)(x_i + x'_i) = zB_i - b_i. \tag{3}$$

This leads to the following result.

LEMMA 2. *The i-conjugate of a lattice point x is always a lattice point in either of the following cases:*

I. $b_i \equiv \alpha_i a_i$, $B_i \equiv \alpha_i A_i$ (α_i*-integer valued*),

II. $A_i \equiv 0$; $a_i \mid b_i$ *and* $a_i \mid B_i$,

where $a \mid b$ *means that* b/a *is integer valued.*

In case I we obtain

$$N(x) = a_i(x_i^2 + \alpha_i x_i) + c_i, \qquad D(x) = A_i(x_i^2 + \alpha_i x_i) + C_i.$$

Thus, if we set $u_i = x_i^2 + \alpha_i x_i$ then $f(x) = F(x_1, \cdots, u_i, \cdots, x_n)$ where F is rational. Thus in case I we consider instead the simpler Diophantine equation

$$(1') \qquad F(x_1, \cdots, u_i, \cdots, x_n) = z,$$

whose solutions obviously include all those obtained from (1) by setting $u_i = x_i^2 + \alpha_i x_i$.

In case II we may first consider the case $B_i \equiv 0$, that is, D independent of x_i. In this case we see that if $f(x)$ represents an integer at all, then it represents all the values of a certain quadratic polynomial, since we may replace x_i by $x_i + mD$ $(m = 0, \pm 1, \cdots)$ to obtain other integers.

In case $B_i \not\equiv 0$ we can write

$$(4) \qquad B_i^2 f = a_i B_i x_i - a_i C_i + B_i b_i + (a_i C_i^2 - b_i B_i C_i + c_i B_i^2)/D.$$

Writing (4) for $x^{(i)}$, subtracting and dividing by $x_i - x'_i$ we obtain

$$(5) \qquad a_i D(x) D(x^{(i)}) = a_i C_i^2 - b_i B_i C_i + c_i B_i^2.$$

4. The method. We can now combine the methods of §§ 2, 3 to state the following.

FINITENESS RESULT. *If every lattice point x satisfying* (1) *has a conjugate lattice point in the exterior of some conical neighborhood of the critical cone, then* (1) *has solutions for at most a finite number of z.*

It will be somewhat more difficult to state a general description of the infinity results and we shall therefore do this by example.

5. The case $n = 2$. We set $x_1 = x$, $x_2 = y$ and have

$$(1'') \qquad f(x, y) = N(x, y)/D(x, y) = z.$$

We first consider the case $\deg D = 4$. That is, $N = ax^2y^2 + \cdots$, $D = Ax^2y^2 + \cdots$, $(A \neq 0)$, where $\cdots$ stands for terms of lower degree. Hence $Af(x,y) = a + N^*/D$ where $\deg N^* \leq 3$. Thus there exist numbers x_0, y_0, M so that if $x \geq x_0$ and $y \geq y_0$ then $f(x,y) \leq M$. We may therefore restrict our attention to the strip neighborhood $|x| < x_0$ or $|y| < y_0$ of the critical cone $xy = 0$.

This reduces the problem to the consideration of a finite number of rational functions of *one* variable

$$g_x(y) = f(x,y), \quad x = 0, \pm 1, \cdots, \pm(x_0 - 1);$$
$$h_y(x) = f(x,y), \quad y = 0, \pm 1, \cdots, \pm(y_0 - 1).$$

The *regular* values of x are those (if any) for which $g_x(y)$ is a polynomial in y; similarly the regular values of y are those for which $h_y(x)$ is a polynomial in x. There are at most 2 regular values for x and for y. If there are such regular values then $g_x(y)$ and $h_y(x)$ may represent an infinity of regular integers. For all other values of (x,y) we obtain at most a finite number of exceptional integers.

We next consider the case $\deg D = 3$. That is, $N = xy(ax + by) + \cdots$, $D = xy(Ax + By) + \cdots$. The method first used can be used with minor modification in the special cases (i) $\deg N = 1$ or (ii) $a/A = b/B$, or (iii) neither N or D contains quadratic terms. Since it involves few new ideas we shall not elaborate on it.

Outside the above cases we shall not be able to proceed without the conjugate point method. If case I of Lemma 2 is satisfied for either x or y then we have seen that by a change of variable we reduce the degree of D to 2, to be discussed below.

For case II of Lemma 2, D must be linear in one of the variables, say y, so that $D = Ax^2y + \cdots$. Hence we may write $Af = a + N^*/D$, $N^* = b^*xy^2 + \cdots$.

The lattice points (x,y) for which $N^* = 0$ lead to at most one integer $z = a/A$. If $N^* \neq 0$ then writing $D = (A_1x^2 + B_1x + C_1)y + A_2x^2 + B_2x + C_2$ we have either $y = 0$ or $A_1y + A_2 = 0$ or there exists an M_1 such that $|x/y| \leq M_1$. If $y = -A_2/A_1$ is an integer then this special value leads to easily determined solutions which may form a regular class or be finite in number. Unless $A_2 = 0$ the value $y = 0$ leads at most to a finite number of exceptional solutions.

Writing now $N^* = a(x)y^2 + b(x)y + c(x)$, $D = B(x)y + C(x)$ we have according to (5)

$$(5') \qquad a(x)(By + C)(By' + C) = aC^2 - bBC + cB^2.$$

The case $a = 0$ can be treated with the method used for $\deg D = 4$. If $a \neq 0$ then the special values of x for which $x = 0$ or $a(x) = 0$ or $B(x) = 0$ lead to easily determined solutions which may contain regular classes. Finally, if the divisibility conditions $a \mid b$ and $a \mid B$ of Lemma 2 are satisfied and $xaB \neq 0$ then we obtain from (5′)

$$\text{(6)} \qquad \frac{y}{x}\cdot\frac{y'}{x} + \frac{C}{xB}\left(\frac{y}{x} + \frac{y'}{x}\right) + \left(\frac{bC}{x^2aB} - \frac{c}{x^2a}\right) = 0$$
$$= \frac{y}{x}\cdot\frac{y'}{x} + \alpha\left(\frac{y}{x} + \frac{y'}{x}\right) + \beta.$$

The degrees of the denominators of α and β are no less than the degrees of their numerators. Hence α, β are bounded. Thus, if we choose $|y| \leqq |y'|$ then there exists an M_2 so that $|y/x| \leqq M_2$.

To sum up. We may obtain regular solutions or a finite number of solutions for $y = -A_2/A_1$, and the values of x for which $x = 0$ or $a(x) = 0$ or $B(x) = 0$. In addition there may be a finite number of exceptional values for $y = 0$. All other lattice points have a conjugate for which both $|x/y| \leqq M_1$ and $|y/x| \leqq M_2$, that is, lying in the exterior of a conical neighborhood of the critical cone $xy = 0$. According to Lemma 1 this leads to at most a finite number of exceptional values of $f(x, y)$.

Finally we consider the case $\deg D = 2$. Let $D_1(x, y) = Ax^2 + Bxy + Cy^2$ be the quadratic part of D and $\Delta = B^2 - 4AC$.

The case $\Delta < 0$, that is D_1 definite, is covered by Lemma 1.

In case $\Delta = 0$ we can make a unimodular transformation so that $D = Ax^2 + B_1x + B_2y + C$. Here there may well be an infinite number of exceptional values. For example if the congruences

$$\text{(7)} \qquad Ax^2 + B_1x + C \equiv \pm 1 \pmod{B_2}$$

have solutions, then for each x satisfying (7) there is a y so that $D(x, y) = \pm 1$ and hence certainly $f(x, y) = z$ is an integer. From the conjugate point method we obtain only the existence of constants M_1, M_2 so that the regular integers are represented by $x = 0$ while all but a finite number of exceptional integers are represented by lattice points satisfying $M_1|x| \leqq |y| \leqq M_2x^2$.

The case $\Delta > 0$, $\Delta \neq$ square, is similar to the preceding one. The equations $D = \pm 1$ may again have infinitely many solutions.

In case $\Delta > 0$, $\Delta =$ square, we can write $D = A(\alpha x + \beta y)(\gamma x + \delta y) + B_1x + B_2y + C$ where $(\alpha, \beta) = (\gamma, \delta) = 1$, $\alpha\delta - \beta\gamma \neq 0$. If we make the transformation $u = \alpha x + \beta y$, $v = \gamma x + \delta y$, then every lattice point (x, y) there corresponds a lattice point (u, v). The converse is not true unless

$\alpha\delta - \beta\gamma = \pm 1$, but for finiteness results we do not need it. We can thus restrict our attention to the case

$$D(x,y) = Axy + B_1x + B_2y + C.$$

If we write $N(x,y) = a_1x^2 + a_2xy + a_3y^2 + b_1x + b_2y + c$, then the divisibility conditions of case II, Lemma 2 become $a_1 \mid a_2$, $a_1 \mid A$, $a_1 \mid B_2$ and $a_3 \mid a_2$, $a_3 \mid A$, $a_3 \mid B_1$ respectively. If both sets of divisibility conditions are satisfied, then from (5) we obtain

$$\begin{aligned}(8)\quad & a_3(Axy + B_1x + B_2y + C)(Axy' + B_1x + B_2y' + C) \\ & = a_3C^2 - (a_2x + b_2)(Ax + B_2)C + (a_1x^2 + b_1x + c)(Ax + B_2)^2\end{aligned}$$

and

$$\begin{aligned}(9)\quad & a_1(Axy + B_1x + B_2y + C)(Ax'y + B_1x' + B_2y + C) \\ & = a_1C^2 - (a_2y + b_1)(Ay + B_1)C + (a_3y^2 + b_2y + c)(Ay + B_1)^2.\end{aligned}$$

From (8) we see that either $x = 0$ or $x = -B_2/A$ or there exists an M_1 such that if we choose $|y| \leqq |y'|$ then $|y/x| \leqq M_1$. Similarly we obtain from (9) that either $y = 0$ or $y = -B_1/A$ or there exists an M_2 such that if we choose $|x| \leqq |x'|$, then $|x/y| \leqq M_2$. Regular values may be represented by $x = -B_2/A$ or $y = -B_1/A$; the values $x = 0$ or $y = 0$ give rise to only a finite number of exceptional values unless they happen to coincide with the regular values. All lattice points (x,y) which are not conjugate to one of the above points have a conjugate—obtained by choosing x minimal and y minimal for that x—which lies in the exterior of a conical neighborhood of the critical cone $xy = 0$. Hence there is only a finite number of exceptional values of z.

6. An example $n = 3$. We consider the equation

$$f(x,y,z) = (x^2 + y^2 + z^2)/(xyz + 1) = u.$$

This is the analogue of [2, Example 1]. The divisibility conditions of Lemma 2, case II are satisfied for x, y, z; and (5) becomes (for the variable z)

$$(xyz + 1)(xyz' + 1) = 1 + x^2y^2(x^2 + y^2).$$

Hence either $xy = 0$ or, if $|z| \leqq |z'|$ then

$$(10)\qquad (xyz + 1)^2 \leqq 1 + x^2y^2(x^2 + y^2).$$

We can obtain the analogous inequality for the other two variables.

Thus the regular values of $f(x, y, z)$ are the ones obtained on the critical cone $xyz = 0$; that is, the numbers which are the sums of two squares.

If $xyz \neq 0$ we may restrict our attention to the points (x, y, z) for which $0 < |x| \leqq y \leqq z$. In case $xyz > 0$ equation (10) yields $z^2 < x^2 + y^2$ and hence $y^2 > z^2/2$. Thus $f(x, y, z) < 3z^2/2^{-\frac{1}{2}}xz^2 = 3 \cdot 2^{\frac{1}{2}}/x$. The only possible positive values for u which can be represented by exceptional lattice points are therefore 1, 2, 3, 4. It is easily seen that the equations $f(1, y, z) = 1$, $f(2, y, z) = 1$, $f(3, y, z) = 1$, $f(4, y, z) = 1$ have no solutions with $yz \neq 0$. The equations $f(1, y, z) = 2$, $f(2, y, z) = 2$ do have solutions with $yz \neq 0$; however these solutions can be seen to be conjugates of solutions with $yz = 0$. Finally $f(1, y, z) = 3$ and $f(1, y, z) = 4$ have no solutions.

We have thus seen that there are no positive exceptional values of u, and that the positive regular values of u are represented only by the conjugates of regular lattice points.

By inspection we obtain the negative exceptional values $f(-1, 2, 2) = -3$, $f(-1, 1, 2) = -6$. From (10) we obtain for $xyz < 0$

$$x^2y^2(z-1)^2 \leqq x^2y^2(1 + x^2 + y^2) \quad \text{or} \quad 3y^2 + 1 \geqq (z-1)^2.$$

Hence

$$f(x, y, z) \leqq 3z^2/[|x| z(z-1)/3^{\frac{1}{2}} - 1] \leqq (3 \cdot 3^{\frac{1}{2}}/|x|) \cdot z/(z-2).$$

We see that for $z \geqq 8$ this leads to $|f| \leqq 6$. Thus we need inspect only a finite number of cases to exclude exceptional values less than -6. For $z \geqq 8$ the values -4, -5 are possible only for $x = -1$ but the equations $f(-1, y, z) = -4$, $f(-1, y, z) = -5$ have no solutions. To exclude $u = -2$ we observe that $f(-1, y, z) = -2$, $f(-2, y, z) = -2$ are both impossible. The impossibility of $u = -1$ follows from the impossibility of the congruence $x^2 + y^2 + z^2 + xyz \equiv -1 \pmod 4$.

7. Infinity results. In [2] we discussed the case

$$f(x, y) = (x^2 + 2y^2)/(1 + xy)$$

which violates the divisibility condition of Lemma 2, case II for y. We can now argue that there must exist an infinity of exceptional values as follows.

We have $x' = -x + yz$, $y' = -y + \frac{1}{2}xz$. For even z this leads to the conjugate point method and our previously discussed finiteness result. For odd z we see that a non-lattice point

$$x = \xi \cdot 2^{-k}, \qquad y = \eta \cdot 2^{-k-\epsilon} \qquad\qquad (\xi, \eta \text{ odd}; \epsilon = 0, 1)$$

has a conjugate which is a lattice point. Hence every integer represented by

$$f_k(x,y) = (x^2 + 2y^2)/(2^k + xy)$$

for odd x, y is also represented by $f(x,y)$. But every f_k represents at least four exceptional integers for odd x, y; namely those obtained from $xy = -2^k \pm 1$. It is clear that these values for f_k are unbounded with k. Hence $f(x,y)$ represents infinitely many odd exceptional numbers. It is easy to apply this process to other cases, but it may be difficult to formulate a general theorem.

8. Conclusion. Our method is clearly not restricted to the classes of equations which we have considered here. Changes in variable may bring a rational function into the form we considered.

A more interesting possibility is that of increasing the number of variables in order to decrease the degree of the equation in each variable and then restrict attention to the case in which the new variables are functions of the old variables.

As an example of this last possibility we discuss the equation

$$u = (x^4 + x^2 + y^2)/(x^3y + 1). \tag{11}$$

The substitution $x^2 = z$ brings this to the form

$$u = (x^2 + y^2 + z^2)/(xyz + 1) \tag{12}$$

which we discussed in Section 6. We see therefore immediately that u must be either the sum of two squares or one of the exceptional values -3, -6. For $x = -1$, $y = 2$ we obtain $u = -6$, while $x = 2$, $y = -1$ yields $u = -3$. Now for the regular values of u we know that they must be represented by a lattice point (x,y,z) so that $z = x^2$ and so that there is a conjugate of one of the forms $(0,a,b)$, $(a,0,b)$, $(a,b,0)$. From (5) we see that

$$x' = -x + uyz, \qquad y' = -y + uxz, \qquad z' = -z + uxy.$$

Thus, if (x_1,y_1,z_1) is a conjugate of (x,y,z), then $x_1 \equiv \pm x$, $y_1 \equiv \pm y$, $z_1 \equiv \pm z \pmod u$. Hence if (x,y,x^2) is conjugate to $(0,a,b)$ then $x \equiv x^2 \equiv 0 \pmod{(a^2+b^2)}$ and $b \equiv \pm x^2 \equiv 0 \pmod{(a^2+b^2)}$ which implies $b = 0$. The regular values thus obtained are therefore exactly those obtained from (11) by setting $x = 0$, that is the squares.

If (x,y,x^2) is conjugate to $(a,0,b)$ then, since $y \equiv 0 \pmod{(a^2+b^2)}$, at every stage, $x \equiv \pm a$, $x^2 \equiv \pm b \pmod{(a^2+b^2)^2}$ and hence $b^2 \equiv a^4 \pmod{(a^2+b^2)^2}$, but $b^2 - a^4 < (a^2+b^2)^2$. Hence $b^2 = a^4$ and $u = a^2 + a^4$

is the regular value obtained from (11) by setting $y=0$. Finally, if (x, y, x^2) is conjugate to $(a, b, 0)$ then

$$x \equiv \pm a \pmod{a^2+b^2}, \; x^2 \equiv a^2 \equiv 0 \pmod{a^2+b^2}.$$

This is possible only if $b=0$ which we have discussed before.

The conjugate point idea is, of course, not restricted to equations which are of second degree in the unknowns, however, our divisibility conditions of Lemma 2 will then have to be replaced by more complicated and cumbersome conditions.

UNIVERSITY OF CALIFORNIA AT LOS ANGELES AND
THE INSTITUTE FOR ADVANCED STUDY.

REFERENCES.

[1] E. S. Barnes, "On the Diophantine equation $x^2+y^2+c=xyz$," *Journal of the London Mathematical Society*, vol. 28 (1953), pp. 242-244.

[2] K. Goldberg, M. Newman, E. G. Straus and J. D. Swift, "The representation of integers by binary quadratic rational forms," *Archiv der Mathematik*, vol. 5 (1954), pp. 12-18.

[3] W. H. Mills, "A system of quadratic Diophantine equations," *Pacific Journal of Mathematics*, vol. 3 (1953), pp. 209-200.

[4] ———, "A method for solving certain Diophantine equations," *Proceedings of the American Mathematical Society*, vol. 5 (1954), pp. 473-475.

IDEALS AND POLYNOMIAL FUNCTIONS.*[1]

By D. J. Lewis.

1. Introduction. Let K be either an algebraic number field or a function field over a finite field or any completion of such fields under a rank one, non-archimedean valuation. Let $\mathfrak{O}$ be the integrally closed ring of integers of K, $\mathfrak{p}$ any prime ideal of $\mathfrak{O}$, and π any element of $\mathfrak{p}$ not in $\mathfrak{p}^2$. Then $\mathfrak{O}/\mathfrak{p}$ is isomorphic to a finite field $GF(q)$, where $q = p^f$, p a rational prime. Let $B_{\mathfrak{p}^m} = B_m$ be the set of polynomials in $\mathfrak{O}[x]$, which when considered as functions on $\mathfrak{O}$ map $\mathfrak{O}$ into $\mathfrak{p}^m$; i.e.,

$$B_m = \{f(x) \text{ in } \mathfrak{O}[x] \text{ such that } f: \mathfrak{O} \to \mathfrak{p}^m\}.$$

Clearly B_m is an ideal in $\mathfrak{O}[x]$.

It is well known [1] that $B_1 = (x^q - x, \pi)$, in fact this result has now become a part of elementary algebra and number theory [2]. Here we analyze B_m, when $m \neq 1$. Clearly $B_0 = \mathfrak{O}[x]$, and $B_m \supset B_{m+1}$. We show that B_m is generated by $m+1$ elements, and give a specific set of generators. The proof is by induction and is of an elementary nature. The results may also be viewed as a concrete realization of a general theory of rings, see [3]; results for the case of polynomials of several indeterminates are also obtained. Because of the computations involved, we first consider the case of one indeterminate and then outline the steps necessary for the case of several indeterminates. These results have been applied to a study of Diophantine equations, which will appear in another paper.

2. Preliminaries. Define

$$\tau_0(x) = x,\ \tau_{n+1}(x) = \tau_n{}^q(x) - \pi^{q^n-1}\tau_n(x) \quad \text{for } n \geqq 0.$$

Let $Q(n) = (q^n - 1)/(q - 1)$. It is easily verified that $\tau_n(x)$ is in $B_{Q(n)}$.

The expression of a rational positive integer m as

$$m = \mu_1 + \mu_2 Q(2) + \cdots + \mu_t Q(t),$$

where $0 \leqq \mu_i \leqq q$, $\mu_t \neq 0$, and $\mu_i = 0$ for $1 \leqq i < j$ if $\mu_j = q$, is unique. We use it to define the following polynomials:

* Received March 29, 1954.

[1] This paper is a part of the results presented to the American Mathematical Society, August 31, 1953, under the title: "Polynomial functions over the residue ring $\mathfrak{O}/\mathfrak{p}^m$." This work was done while the author was a National Science Foundation Fellow at the Institute for Advanced Study.

$$\lambda_0 = 1,\ \lambda_m = \tau_1^{\mu_1}\tau_2^{\mu_2}\cdots\tau_t^{\mu_t} \text{ if } m > 0.$$

Clearly λ_m is an element of B_m.

Let $A_m = (\lambda_m, \pi\lambda_{m-1}, \cdots, \pi^{m-1}\lambda_1, \pi^m)$, $m \geqq 1$. Then $A_m \subset B_m$. We contend that $B_m = A_m$. As we shall sometime make a change in variable, for clarity we may sometimes write $A_m(x)$ or $B_m(x)$ to indicate we are operating in $\mathfrak{O}[x]$.

We make several observations concerning A_m.

THEOREM I. i) $A_m \supset A_{m+1}$.

ii) $A_m \cdot A_n \subset A_{m+n}$.

iii) *If* $g(z)$ *is in* $\mathfrak{O}[x]$ *and* $f(x)$ *is in* $A_m(x)$, *then* $f(g(z))$ *is in* $A_m(z)$; *symbolically* $A_m(g(z)) \subset A_m(z)$.

Proof. In view of the definition of A_m, (i) is evident, provided λ_{m+1} is in A_m. Consider the expansion

$$m + 1 = \nu_1 + \nu_2 Q(2) + \cdots + \nu_t Q(t).$$

If $\nu_1 \neq 0$, $\lambda_{m+1} = \tau_1\lambda_m$ and hence is in A_m. If $\nu_1 = \nu_2 = \cdots = \nu_r = 0$, and $\nu_{r+1} \neq 0$, then $\lambda_{m+1} = \lambda_m - \pi^{q^r-1}\lambda_{(m-q^r+1)}$, which is in A_m.

By definition τ_n^α is in $A_{\alpha Q(n)}$ if $0 \leqq \alpha \leqq q$, but because of (i) we have this for all $\alpha \geqq 0$. Then by induction $\tau_n^\beta\lambda_m$ is in $A_{m+\beta Q(n)}$ if $q \geqq \beta \geqq 0$, and finally $\pi^s\lambda_m\lambda_n$ is in A_{s+m+n}. Thus proving (ii).

Using (ii) and the definition of τ_i and λ_s, we obtain $\lambda_s^q \equiv \pi^{qs-s}\lambda_s \pmod{A_{qs+1}}$.

Observe that the binomial coefficients $C_i^q \equiv 0 \pmod{\mathfrak{p}}$ if $1 \leqq i \leqq q-1$, also that $\tau_1(z^r) = \sum_{i=1}^{n} z^{(r-i)q+i-1}\tau_1(z)$. It follows immediately that $\tau_1(g(z))$ is in $A_1(z)$. Because of (ii) it is clearly sufficient for proving (iii) to show that $\tau_n(g(z))$ is in $A_{Q(n)}(z)$. This is proved by induction:

Suppose $\tau_k(g(z))$ is in $A_{Q(k)}(z)$ if $1 \leqq k \leqq n$. Then:

$$\begin{aligned}
\tau_{n+1}(g(z)) &= \Big[\sum_{s=0}^{Q(n)} h_s(z)\pi^{Q(n)-s}\lambda_s(z)\Big]^q - \pi^{q^n-1}\sum_{s=0}^{Q(n)} h_s(z)\pi^{Q(n)-s}\lambda_s(z)\\
&\equiv \sum_{s=0}^{Q(n)} \big[h_s(z^q)\pi^{qQ(n)-qs}\lambda_s^q(z) - h_s(z)\pi^{Q(n+1)-s-1}\lambda_s(z)\big] \pmod{A_{Q(n+1)}}\\
&\equiv \sum_{s=0}^{Q(n)} \big[h_s(z^q)\pi^{qQ(n)-s}\lambda_s(z) - h_s(z)\pi^{qQ(n)-s}\lambda_s(z)\big] \pmod{A_{Q(n+1)}}\\
&\equiv \sum_{s=0}^{Q(n)} \pi^{qQ(n)-s}\lambda_s(z)\big[h_s(z^q) - h_s(z)\big] \pmod{A_{Q(n+1)}}\\
&\equiv \sum_{s=0}^{Q(n)} \pi^{qQ(n)-s}\lambda_s(z)\tau_1(h(z)) \equiv 0 \pmod{A_{Q(n+1)}}.
\end{aligned}$$

If $m = \mu_1 + \mu_2 Q(2) + \cdots + \mu_t Q(t)$, define

$$m^* = \mu_2 + \mu_3 Q(2) + \cdots + \mu_t Q(t-1).$$

Then $m - m^* = \sum_{i=1}^{t} \mu_i q^{i-1}$ and $m \leqq q$ implies $m^* = 0$.

We shall have use of the following lemma.

LEMMA 1. *For every a in $\mathfrak{O}$, there exists a c in $\mathfrak{O}$ such that $a \equiv c \pmod{\mathfrak{p}}$ and such that for every λ_s there is a polynomial $h_s(z)$ in B_{s^*-1} such that*

$$\lambda_s(c - \pi z) = \pi^{s-s^*} z^{\mu_1(s)} \lambda_{s^*}(z) + \pi^{s-s^*+1} h_s(z).$$

Proof. Take $c = a + \tau_1(a)$. Then $\tau_1(c) \equiv 0 \pmod{\mathfrak{p}^2}$ and consequently $\tau_1(c - \pi z) = \pi z + \pi^2 h_1(z)$. By induction we show that

$$\tau_{n+1}(c - \pi z) \equiv \pi^{q^n} \tau_n(z) \qquad (\operatorname{mod} \mathfrak{p}^{q^n+1}),$$

the congruence being coefficient-wise. Suppose such is the case for τ_k, where $1 \leqq k \leqq n$. Then

$$\begin{aligned}
\tau_{n+1}(c - \pi z) &= \tau_n{}^q(c - \pi z) - \pi^{q^{n-1}} \tau_n(c - \pi z) \\
&= \pi^{q^n}[\tau_{n-1}(z) + \pi k(z)]^q - \pi^{q^n + q^{n-1} - 1} - 1[\tau_{n-1}(z) + \pi k(z)] \\
&\equiv \pi^{q^n}[\tau_{n-1}{}^q(z) - \pi^{q^{n-1}-1} \tau_{n-1}(z)] \quad (\operatorname{mod} \mathfrak{p}^{q^n+1}) \\
&\equiv \pi^{q^n} \tau_n(z) \quad (\operatorname{mod} \mathfrak{p}^{q^n+1}).
\end{aligned}$$

If $s = \mu_1(s) + \mu_2(s) Q(2) + \cdots + \mu_t(s) Q(t)$, we obtain

$$\begin{aligned}
\lambda_s(c - \pi z) &= \prod_{i=1}^{t} \tau_i{}^{\mu_i(s)}(c - \pi z) = \pi^{s-s^*} \prod_{i=1}^{t} [\tau_{i-1}(z) + \pi k_i(z)]^{\mu_i(s)} \\
&= \pi^{s-s^*} z^{\mu_1(s)} \lambda_{s^*}(z) + \pi^{s-s^*+1} h_s(z).
\end{aligned}$$

Since λ_s and $\pi^{s-s^*} \lambda_{s^*}$ are in B_s, it follows that h_s is in B_{s^*-1}.

3. Proof of the contention that $B_m = A_m$. Let

$$\Gamma = \{\lambda_m \text{ such that } m = \mu_1 + \mu_2 Q(2) + \cdots + \mu_t Q(t)\,;\ 0 \leqq \mu_i < q\}.$$

Then for each non-negative integer r, Γ contains a unique polynomial of degree rq.

Let $f(x)$ be any polynomial in $\mathfrak{O}[x]$. Let d denote the degree of $f(x)$ and define e such that $eq \leqq d < q(e+1)$. Then $f(x)$ can be expressed uniquely as $f(x) = g(x)\lambda(x) + f^*(x)$, where the degree of $g(x)$ is less

than q, $\lambda(x)$ is the unique polynomial in Γ of degree eq and the degree of $f^*(x)$ is less than eq.

If c is the largest power of π dividing any coefficient of $g(x)$ we may write $g(x)=\sum_{i=0}^{c}\pi^i g_i(x)$, where the non-zero coefficients of the $g_i(x)$ are not divisible by π. Clearly the $g_i(x)$ are uniquely determined by $g(x)$, hence by $f(x)$. Continuing the process on the residual polynomial we arrive at a unique expression for $f(x)$ of the form.

$$f(x)=\sum_{\Gamma}\sum_{t} g_{k,t}(x)\pi^t\lambda_k(x), \tag{1}$$

where the $\lambda_k(x)$ are in Γ, the degree of each $g_{k,t}(x)$ is less than q, the non-zero coefficients of the $g_{k,t}(x)$ are not in $\mathfrak{p}$; and almost all of the $g_{k,t}(x)$ are the zero polynomial.

As previously remarked $B_1=A_1$. Assume $B_i=A_i$, for $1\leqq i\leqq m$ and suppose $f(x)$ is in B_{m+1}. Now $B_{m+1}\subset B_m=A_m$, hence using (1) we obtain

$$f(x)\equiv\sum_{s=0}^{m}{}' g_s(x)\pi^{m-s}\lambda_s(x) \pmod{A_{m+1}} \tag{2}$$

where the $g_s(x)$ are of degree less than q and their non-zero coefficients are not in $\mathfrak{p}$, and where the accent mark indicates that the sum ranges over the λ_s, $0\leqq s\leqq m$, which are in Γ.

If $s\leqq m$, we have $s^*<m$, then in light of the induction hypothesis for $s\leqq m$, the $h_s(z)$ of Lemma 1 is in A_{s^*-1}. Hence

$$\begin{aligned} f(c-\pi z)&\equiv\sum_{s=0}^{m}{}' (g_s(a)+\pi w_s(z))\pi^{m-s}\lambda_s(c-\pi z) \pmod{A_{m+1}}\\ &\equiv\sum_{s=0}^{m}{}' g_s(a)[z^{\mu_1(s)}\pi^{m-s^*}\lambda_{s^*}(z)+\pi^{m-s^*+1}h_s(z)] \pmod{A_{m+1}}\\ &\equiv[\textstyle\sum'' \gamma_s z^{\mu_1(s)}]\pi^{m-m^*}\lambda_{m^*}(z)+\pi^{m-m^*+1}R(z) \pmod{A_{m+1}}\end{aligned}$$

where $R(z)$ is a polynomial in A_{m^*-1}; where the double accent indicates that the sum ranges over those s for which λ_s is in Γ and for which $s^*=m^*$; and where $\gamma_s=g_s(a)$ if $g_s(a)$ not in $\mathfrak{p}$ and $\gamma_s=0$ if $g_s(a)$ is in $\mathfrak{p}$.

If $s^*=m^*=t^*$ and $s\neq t$, then $\mu_1(s)\neq\mu_1(t)$, thus the γ_s are coefficients of different powers of z. Since $\mu_1(s)<q$ for all s for which λ_s is in Γ, $G(z)=\sum''\gamma_s z^{\mu_1(s)}$ is a g-polynomial of the type specified in (1).

Let $H(z)=G(z)\lambda_{m^*}(z)+\pi R(z)$, then $f(c-\pi z)\equiv\pi^{m-m^*}H(z) \pmod{A_{m+1}}$. By assumption $f(x)$ is in B_{m+1}, hence $H(z)$ is in B_{m^*+1}. But $m^*+1\leqq m$, thus by the induction hypothesis $B_{m^*+1}=A_{m^*+1}$. Hence, $H(z)$ is in A_{m^*+1},

consequently $G(z)$ is the zero polynomial and we have $g_s(a) \equiv 0 \pmod{p}$, for all s for which λ_s is in Γ and for which $s^* = m^*$.

Since this computation is true for every a in $\mathfrak{O}$, we have that for these s, $g_s(x)$ is in $B_1(x) = A_1(x)$ and hence are the zero-polynomial. In particular if λ_m is in Γ, $g_m(x)$ is the zero polynomial. Using this fact and (2), we obtain

$$f(x) \equiv \pi[\sum_{s=0}^{m-1}{}' g_s(x)\pi^{m-1-s}\lambda_s(x)] = \pi F(x) \pmod{A_{m+1}}.$$

But then $F(x)$ is in $B_m = A_m$ and so must be the zero polynomial. Thus $f(x)$ is in A_{m+1}. Since $f(x)$ was any polynomial from B_{m+1}, we have proved that $B_{m+1} = A_{m+1}$, proving

THEOREM II. *$A_m = B_m$, for all $m \geqq 1$.*

We also obtain that $B_\infty = \bigcap_{m \geq 1} B_m = 0$. For suppose $f(x)$ is in B_∞, let d denote the degree of $f(x)$ and let r be the highest power of π dividing all of the coefficients of $f(x)$. Then $f(x)$ is not in B_{d+r+1}, hence not in B_∞.

4. The case of more than one indeterminant. Where convenient we shall denote a polynomial of $\mathfrak{O}[x_1, x_2, \cdots, x_n]$ by $f(X)$ and sometimes just by f. Let $\mathfrak{V}$ be the n-dimensional vector space over $\mathfrak{O}$. Let

$$\mathfrak{B}_m = \{f \text{ in } \mathfrak{O}[x_1, x_2, \cdots, x_n] \text{ such that } f: \mathfrak{V} \to \mathfrak{p}^m\}.$$

It is well known [1] that $\mathfrak{B}_1 = (\tau_1(x_1), \tau_1(x_2), \cdots, \tau_1(x_n), \pi)$.

Consider all partitions, $(\rho) = (\rho_1, \rho_2, \rho_3, \cdots, \rho_n)$ of m into n non-negative integers; i.e., $m = \rho_1 + \rho_2 + \cdots + \rho_n$, where $\rho_i \geqq 0$. Define

$$\Lambda_m^{(\rho)}(X) = \prod_{i=1}^{n} \lambda_{\rho_i}(x_i).$$

Clearly each $\Lambda_m^{(\rho)}$ is in $\mathfrak{B}_m$. For completeness we shall outline the proof of the following result:

THEROEM III. *$\mathfrak{B}_m = (\pi\mathfrak{B}_{m-1}, \Lambda_m^{(\rho)})$, where (ρ) ranges over all partitions of m.*

Let $\mathfrak{A}_m = (\pi\mathfrak{A}_{m-1}, \Lambda_m^{(\rho)})$, where (ρ) ranges over all partitions of m. Then we obtain

LEMMA 2. i) $\mathfrak{A}_m \supset \mathfrak{A}_{m+1}$

ii) $\mathfrak{A}_m \cdot \mathfrak{A}_n \subset \mathfrak{A}_{m+n}$

iii) *If $g_i(Z)$ are in $\mathfrak{O}[z_1, z_2, \cdots, z_n]$ and $f(X)$ is in $\mathfrak{A}_m(X)$, then $f(g_1(Z), g_2(Z), \cdots, g_n(Z))$ is in $\mathfrak{A}_m(Z)$.*

If $(\rho) = (\rho_1, \rho_2, \cdots, \rho_n)$ define $\rho^* = \sum \rho_i^*$ where ρ_i^* is defined as in Section 2. Then $\rho^* < m$ and $(\rho^*) = (\rho_1^*, \rho_2^*, \cdots, \rho_n^*)$ is a partition of ρ^*. Using Lemmas 1, 2 and definitions, we obtain

LEMMA 3. *For every vector* $\mathfrak{a} = (a_1, a_2, \cdots, a_n)$ *in* $\mathfrak{B}$ *there exists a vector* $\mathfrak{c} = (c_1, c_2, \cdots, c_n)$ *in* $\mathfrak{B}$ *such that* $a_i \equiv c_i \pmod{\mathfrak{p}}$ *for* $1 \leqq i \leqq n$, *and such that for every* $\Lambda_s^{(\sigma)}$ *there is a polynomial* $H^{(\sigma)}$ *in* $\mathfrak{B}_{\sigma^*-1}$ *such that*

$$\Lambda_s^{(\sigma)}(\mathfrak{c} - \pi Z) = \pi^{s-\sigma^*} k^{(\sigma)}(Z) \Lambda_{\sigma^*}^{(\sigma^*)}(Z) + \pi^{s-\sigma^*+1} H^{(\sigma)}, \text{ where } k^{(\sigma)}(Z) = \prod_{i=1}^{n} z_i^{\mu_1(c_i)}.$$

Let Γ_0 be the subset of the $\Lambda_m^{(\rho)}$ which do not have a factor $\tau_i(x_j)$, appearing to the q-th power. We can develop an algorithm which leads to a unique expression for each f in $\mathfrak{D}[x_1, x_2, \cdots, x_n]$ in the form

$$f(X) = \sum_{\Gamma_0} \sum_{t} g_{s,(\sigma),t}(X) \pi^t \Lambda_s^{(\sigma)}(X) \tag{3}$$

where the $\Lambda_s^{(\sigma)}$ are in Γ_0, the degree of each indeterminate in $g_{s,(\sigma),t}$ is less than q, the coefficients of the $g_{s,(\sigma),t}$ are either zero or not in $\mathfrak{p}$, and almost all of the $g_{s,(\sigma),t}$ are the zero polynomial.

As noted $\mathfrak{B}_1 = \mathfrak{A}_1$, we suppose $\mathfrak{B}_i = \mathfrak{A}_i$ for $i \leq m$ and prove $\mathfrak{B}_{m+1} = \mathfrak{A}_{m+1}$. Let f be in $\mathfrak{B}_{m+1} \subset \mathfrak{B}_m = \mathfrak{A}_m$. Then

$$f(X) = \sum_{s=0}^{m} \sum_{(\sigma)}{}' \pi^{m-s} g_{s,(\sigma)}(X) \Lambda_s^{(\sigma)}(X) \pmod{\mathfrak{A}_{m+1}}$$

where the accent indicates the sum ranges over that $\Lambda_s^{(\sigma)}$ in Γ_0, and the $g_{s,(\sigma)}$ are as in (3).

Let $u = \operatorname{Max}\{\sigma^*$, where (σ) is a partition of s and $s \leqq m\}$. Then

$$f(X) = \sum_{r=0}^{u} f_r(X) \pmod{\mathfrak{A}_{m+1}}, \text{ where } f_r(X) = \sum_{(\sigma)}^{\sigma^*=r}{}' \pi^{m-s} g_{s,(\sigma)}(X) \Lambda_s^{(\sigma)}(X).$$

Here the sum ranges over those integers s and their partitions (σ) for which $\Lambda_s^{(\sigma)}$ is in Γ_0, $s \leqq m$ and $\sigma^* = r$. Then

$$f(\mathfrak{c} - \pi Z) \equiv \pi^{m-u} G(Z) + \pi^{m-u+1} R(Z) \pmod{\mathfrak{A}_{m+1}},$$

where $f_u(\mathfrak{c} - \pi Z) \equiv \pi^{m-u} G(Z) \pmod{p^{m-u+1}}$ (the congruence being coefficient-wise), where

$$G(Z) = \sum K^{(\omega)}(Z) \Lambda_u^{(\omega)}(Z)$$

where the sum ranges over partitions (ω) of u. And $K^{(\omega)}(Z) = \sum \gamma_{s,(\sigma)} k^{(\sigma)}(Z)$, where the sum ranges over those integers $s \leqq m$ and their partitions (σ) for which $(\sigma^*) = (\omega)$ and for which $\Lambda_s^{(\sigma)}$ is in Γ_0. The $\gamma_{s,(\sigma)} = g_{s,(\sigma)}(\mathfrak{a})$, if $g_{s,(\sigma)}(\mathfrak{a})$ is not in $\mathfrak{p}$ and $\gamma_{s,(\sigma)} = 0$ if $g_{s,(\sigma)}(\mathfrak{a})$ is in $\mathfrak{p}$.

If (σ) and (δ) are either different partitions of the same integer or are partitions of different integers, there is a j such that $\sigma_j \neq \delta_j$. If $(\sigma^*) = (\delta^*)$, then $\mu_1(\sigma_j) \neq \mu_1(\delta_j)$. Hence in $K^{(\omega)}$ the various $\gamma_{s,(\sigma)}$ are coefficients of different monomials. Since $\Lambda_s^{(\sigma)}$ were in Γ_0, the degree of each indeterminate in the $k^{(\sigma)}$, and hence in the $K^{(\omega)}$, is less than q.

Since $f(\mathfrak{c} - \pi Z)$ is in B_{m+1}, $G(Z) + \pi R(Z)$ must be in B_{u+1}. But $u + 1 \leqq m$, hence $G(Z) + \pi R(Z)$ is in A_{u+1}. Consequently $G(Z)$ is the zero polynomial, and thus the $K^{(\omega)}(Z)$ are the zero polynomial. Hence for those integers s and their partitions (σ) for which $\Lambda_s^{(\sigma)}$ is in Γ_0, $s \leqq m$ and $\sigma^* = u$, we have $g_{s,(\sigma)}(\mathfrak{a}) \equiv 0 \pmod{p}$. Since this is true for every $\mathfrak{a}$ in $\mathfrak{B}$, we must have for these s and (σ) that $g_{s,(\sigma)}$ is the zero polynomial. Consequently $f_u(X)$ is the zero polynomial. One can now continue step-wise, and show that for each r, $f_r(X)$ is the zero polynomial. Hence $f(X)$ is in $\mathfrak{A}_{m+1}$.

5. Generalization. If we let $\mathfrak{O}$ be a ring of algebraic integers and let $\mathfrak{m}$ be an ideal in $\mathfrak{O}$ which is not a power of a prime ideal in $\mathfrak{O}$ and consider $B_{\mathfrak{m}}$, or $\mathfrak{B}_{\mathfrak{m}}$, we have considerable more difficulty. If $\mathfrak{m} = \mathfrak{p}_1^a \mathfrak{p}_2^b \cdots \mathfrak{p}_s^c$, $B_{\mathfrak{m}} \neq B_{\mathfrak{p}_1^a} B_{\mathfrak{p}_2^b} \cdots B_{\mathfrak{p}_s^c}$. For let Z be the ring of rational integers and let p and q be primes in Z, then $x(x^{p-1} - 1)(x^{q-1} - 1)$ is in B_{pq}. This illustrates the type of element in $B_{\mathfrak{m}}$. We leave further discussion of $B_{\mathfrak{m}}$ to another time.

UNIVERSITY OF NOTRE DAME.

REFERENCES.

[1] C. Chevalley, "Demonstration d'une hypothese de M. Artin," *Abhandlungen aus dem Mathematischen Seminar der Hamburgischen Universitat*, vol. 11 (1935), pp. 73-75.

[2] H. Davenport, *Higher Arithmetic*, Hutchinson's University Library, 1952.

[3] E. Snapper, "Completely Primary Rings. IV," *Annals of Mathematics*, vol. 55 (1952), pp. 46-64.

A GENERAL THEORY OF ALGEBRAIC GEOMETRY OVER DEDEKIND DOMAINS, I.*

The Notion of Models.

By MASAYOSHI NAGATA.

In the present sequence of papers, we want to study a general theory of algebraic geometry over a ring, which is a field or a Dedekind domain, under the restriction that the almost finite integral extensions of this ring are finite. (Observe that this condition is satisfied by fields, by complete discrete valuation rings and by Dedekind domain of characteristic zero.)

The writer wishes at first to express his hearty thanks to Professor C. Chevalley, to whose lectures[1] at Kyôto University the writer owes many ideas and who gave the writer many suggestions during the preparation of the present paper.

In Chapter 1, we prove some preliminary results on rings (mainly on spots, which will play an important rôle in our study). In Chapter 2, we study the notion of models of function fields.

In Chapter 1 we first prove the normalization theorem in a generalized form (§ 1) and then we define the notion of spots and study some of their properties; here the notions of affine rings and of function fields are also defined (§§ 2-4). Applying a result in § 4, we prove the finiteness of the derived normal ring of an affine ring in § 5. In § 6, we prove some lemmas on valuation rings.

In Chapter 2, we first introduce the notions of places and of models (§§ 1-2) and then we introduce the notion of specializations (§ 3). Then we study the notion of joins of models (§ 4) and prove the existence of the derived normal model of a model (§ 5). In §§ 6-7, we introduce the Zariski topology on models and in § 8 we introduce the notions of induced model, local model and reduced model. In § 9, we show that under a certain restriction on the function field under consideration, the notion of model is equivalent to the notion of abstract variety in the sense of Weil [12].

* Received May 16, 1955; revised November 10, 1955.

[1] Professor C. Chevalley lectured at Kyôto University in January of 1954. Our definition of models is an adaptation of his to our case. Main results in Chapter 2 of the present paper were shown by him for the case of algebraic geometry over field in his lecture.

Terminology and notations.

Besides the terminology which was used in Nagata [10], we use the following terms: A ring is called *quasi-local* if it has only one maximal ideal. When $\mathfrak{o}$ and $\mathfrak{o}'$ are quasi-local rings, we say that $\mathfrak{o}'$ *dominates* $\mathfrak{o}$ if $\mathfrak{o}$ is a subring of $\mathfrak{o}'$ and if the maximal ideal of $\mathfrak{o}'$ lies over that of $\mathfrak{o}$ (we say that an ideal $\mathfrak{a}'$ of a ring $\mathfrak{o}'$ lies over an ideal $\mathfrak{a}$ of its subring $\mathfrak{o}$ if $\mathfrak{a} = \mathfrak{a}' \cap \mathfrak{o}$). Observe here that domination defines a partial order.

A ring $\mathfrak{o}$ is called a *semi-local* ring if it has only a finite number of maximal ideals and if the topology of $\mathfrak{o}$ introduced by taking all powers of its J-radical as a system of neighborhoods of zero is Hausdorff; a quasi-local semi-local ring is a *local* ring. But since we treat mainly the Noetherian case, we shall mean by local or semi-local ring a *Noetherian* local or semi-local ring, unless the contrary is explicitly stated.

When $\mathfrak{o}$ is an integral domain, the integral closure of $\mathfrak{o}$ in its field of quotients is called the *derived normal ring* of $\mathfrak{o}$. Let $\mathfrak{o}$ be an integral domain. An integral extension $\mathfrak{o}'$ of $\mathfrak{o}$ is said to be *almost finite* if the field of quotients of $\mathfrak{o}'$ is a finite algebraic extension of that of $\mathfrak{o}$ (see [10]); we say that $\mathfrak{o}$ satisfies the *finiteness condition for integral extensions* if every almost finite integral extension of $\mathfrak{o}$ is a finite $\mathfrak{o}$-module. Observe that, if $\mathfrak{o}$ satisfies the finiteness condition for integral extensions, then so does any ring of quotients of $\mathfrak{o}$.

Let $\mathfrak{o}$ be a ring and let S be a set of elements of a ring containing $\mathfrak{o}$. Consider the ring $\mathfrak{o}' = \mathfrak{o}[S]$. Let T be the intersection of the complements of all prime divisors of ideals of $\mathfrak{o}'$ generated by maximal ideals of $\mathfrak{o}$. Then the ring $\mathfrak{o}'_T$ is denoted by $\mathfrak{o}(S)$. Observe that if S is a set of independent elements over a local ring $\mathfrak{o}$ with maximal ideal $\mathfrak{m}$, then $\mathfrak{o}(S) = \mathfrak{o}[S]_{\mathfrak{m}\mathfrak{o}[S]}$. (We shall use mainly this last case.)

Though the notion of rank of rings was defined in [10], we shall repeat it again: A ring $\mathfrak{o}$ is said to be of *rank* n if there exists a chain $\mathfrak{p}_0 \subset \mathfrak{p}_1 \subset \cdots \subset \mathfrak{p}_n$ of $n+1$ prime ideals $\mathfrak{p}_i$ of $\mathfrak{o}$ and if there exists no such chain with more prime ideals; here the symbol $\subset$ means "included in and different from" (to indicate only "included in," we shall use the symbol $\subseteq$) and prime ideals mean those which are different from the ring. When $\mathfrak{a}$ is an ideal of $\mathfrak{o}$, rank $\mathfrak{o}/\mathfrak{a}$ is called the *co-rank* of $\mathfrak{a}$. The rank of a prime ideal $\mathfrak{p}$ of $\mathfrak{o}$ is defined as the rank of $\mathfrak{o}_{\mathfrak{p}}$; the *rank* of an ideal $\mathfrak{a}$ is the minimum of the ranks of the prime divisors of $\mathfrak{a}$.

Results assumed to be known.

Besides elementary results on fields and rings of polynomials, we need

some results on commutative rings: (1) For the general theory of commutative rings, results which are contained in Nagata [10] are assumed to be known. (2) For the theory of local rings, we assume that the following lemmas are known:

LEMMA 0.1. *If* $\mathfrak{o}$ *is a local ring with maximal ideal* $\mathfrak{m}$, *then the completion* $\mathfrak{o}^*$ *of* $\mathfrak{o}$ *is a local ring with maximal ideal* $\mathfrak{m}\mathfrak{o}^*$. *Further, rank* $\mathfrak{o} =$ *rank* $\mathfrak{o}^*$ *and if* $\mathfrak{a}$ *is an ideal of* $\mathfrak{o}$, *then* $\mathfrak{a}\mathfrak{o}^* \cap \mathfrak{o} = \mathfrak{a}$. (See Krull [5], Cohen [2], Nagata [9], Samuel [11].)

LEMMA 0.2. *Assume that* $\mathfrak{o}$ *is a semi-local ring with maximal ideals* $\mathfrak{p}_1, \cdots, \mathfrak{p}_h$. *Then the completion of* $\mathfrak{o}$ *is the direct sum of the completions of the local rings* $\mathfrak{o}_{\mathfrak{p}_1}, \cdots, \mathfrak{o}_{\mathfrak{p}_h}$. (See Chevalley [1], Nagata [9], Samuel [11].)

(A proof can be given as follows: Set $\mathfrak{m} = \bigcap_i \mathfrak{p}_i$. Then the completion $\mathfrak{o}^*$ of $\mathfrak{o}$ is the limit space of the inverse system $\{\mathfrak{o}/\mathfrak{m}^n; n = 1, 2, \cdots\}$. Since the $\mathfrak{p}_i$'s are maximal ideals, $\mathfrak{o}/\mathfrak{m}^n$ is isomorphic to the direct sum of rings $\mathfrak{o}/\mathfrak{p}_i^n$ $(1 \leqq i \leqq h)$. But $\mathfrak{o}/\mathfrak{p}_i^n = \mathfrak{o}_{\mathfrak{p}_i}/\mathfrak{p}_i^n\mathfrak{o}_{\mathfrak{p}_i}$. Therefore $\mathfrak{o}^*$ is isomorphic to the direct product (sum) of the limit spaces of the inverse systems $\{\mathfrak{o}_{\mathfrak{p}_i}/\mathfrak{p}_i^n\mathfrak{o}_{\mathfrak{p}_i}\}$, whose limits coincide with the completions of the $\mathfrak{o}_{\mathfrak{p}_i}$.)

LEMMA 0.3. *Let* $\mathfrak{a}$ *be an ideal of a semi-local ring* $\mathfrak{o}$ *and let* $\mathfrak{o}^*$ *be the completion of* $\mathfrak{o}$. *Then* $\mathfrak{a}\mathfrak{o}^* \cap \mathfrak{o} = \mathfrak{a}$ and $\mathfrak{o}^*/\mathfrak{a}\mathfrak{o}^*$ *is the completion of* $\mathfrak{o}/\mathfrak{a}$. (See Chevalley [1], Nagata [9], Samuel [11].)

LEMMA 0.4. *Let* $\mathfrak{o}$, $\mathfrak{o}^*$ *and* $\mathfrak{a}$ *be as in Lemma* 0.3. *If* b *is an element of* $\mathfrak{o}$, *then* $\mathfrak{a}\mathfrak{o}^* : b\mathfrak{o}^* = (\mathfrak{a} : b\mathfrak{o})\mathfrak{o}^*$. (See Zariski [15], Nagata [9], Samuel [11].)

COROLLARY. *If an element* a *of* $\mathfrak{o}$ *is not a zero-divisor in* $\mathfrak{o}$, *then* a *is not a zero-divisor in* $\mathfrak{o}^*$. (Chevalley [1])

LEMMA 0.5. *Let* $\mathfrak{o}$ *be a complete local ring with maximal ideal* $\mathfrak{m}$. *Assume that a local ring* $\mathfrak{o}'$ *dominates* $\mathfrak{o}$. *If* $\mathfrak{o}'/\mathfrak{m}\mathfrak{o}'$ *is a finite* $\mathfrak{o}/\mathfrak{m}$*-module, then* $\mathfrak{o}'$ *is a finite* $\mathfrak{o}$*-module and is a complete local ring.* (See Chevalley [1], Cohen [2], Nagata [9], Samuel [11].)

LEMMA 0.6. *Let* $\mathfrak{o}$ *and* $\mathfrak{o}'$ *be semi-local rings which satisfy the following conditions*: 1) $\mathfrak{o}$ *is a subring of* $\mathfrak{o}'$, 2) $\mathfrak{o}'$ *is a finite* $\mathfrak{o}$*-module generated by* $y_1, \cdots, y_n$ *and* 3) *every nonzero element of* $\mathfrak{o}$ *is not a zero-divisor in* $\mathfrak{o}'$. *Then* I) $\mathfrak{o}$ *is a subspace of* $\mathfrak{o}'$, II) *the completion* $\mathfrak{o}'^*$ *of* $\mathfrak{o}'$ *is the module generated by* $y_1, \cdots, y_n$ *over the completion* $\mathfrak{o}^*$ *of* $\mathfrak{o}$, III) *if elements* $x_1, \cdots, x_r$ *of* $\mathfrak{o}'$ *are linearly independent over* $\mathfrak{o}$, *then they are linearly*

independent over $\mathfrak{o}^*$ *and* IV) *if an element* a *of* $\mathfrak{o}^*$ *is not a zero-divisor in* $\mathfrak{o}^*$, *then it is not a zero-divisor in* $\mathfrak{o}'^*$. (See Chevalley [1], Nagata [9], Samuel [11].)

LEMMA 0.7. *A regular local ring is a normal ring.* (See Krull [5], Cohen [2], Nagata [9], Samuel [11].)

LEMMA 0.8. *Let* $\mathfrak{o}$ *be a complete local ring. Assume that* $\mathfrak{o}$ *contains a field or dominates a descrete valuation ring with prime element* p; *in the latter case we asume further that* $p\mathfrak{o}$ *is of rank* 1. *Then there exists an unramified regular local ring* $\mathfrak{r}$ *contained in* $\mathfrak{o}$ *such that* $\mathfrak{o}$ *is a finite* $\mathfrak{r}$*-module.* (See Cohen [2], Samuel [11].) (It was communicated to the writer that a much simplified proof of this lemma is given by Mr. Narita in a forthcoming paper. (Added October, 1955.))

LEMMA 0.9. *If* $\mathfrak{r}$ *is a complete, unramified regular local ring, then any prime ideal of rank* 1 *in* $\mathfrak{r}$ *is principal.* (See Cohen [2], Nagata [9], Samuel [11].)

Remark. This last lemma holds without the assumption that $\mathfrak{r}$ is complete (see Nagata [9]; the proof will be repeated in the second paper of this sequence), as was announced by Mr. Y. Mori in the spring meeting of the Mathematical Society of Japan in 1949.

Numbering.

Numbering of lemmas will begin anew in each section; numbering of propositions and theorems will begin in each chapter. When we refer to a lemma in another section, we shall use notation such as Lemma 1.2.3, the first number, the second one and the third one indicating the number of chapter, section and the lemma respectively; as for the theorems or propositions, we shall use notations like Theorem 1.1, the first number indicating the number of the chapter.

On the restriction of ground rings.

As was stated above, ground rings are assumed to satisfy the finiteness condition for integral extensions. But the definition of spots, affine rings, function fields, models and so on may be given without making use of this condition. When we want to talk about these notions without assuming that the ground ring satisfies the finiteness condition for integral extensions, we shall say: "in the non-restricted case."

Chapter 1. Preliminaries from the Theory of Rihgs.

Most of the results in the present chapter are not new: Results in § 1 are essentially contained in Nagata [7], [9]. The main results in § 4 is a slight generalization of a result in Nagata [8]. The results in § 6 are well known, but they are basic for the theory of valuation rings.

1. Normalization theorem.

LEMMA 1. *Let $\boldsymbol{k}$ be a field and let $x_1, \cdots, x_n$ be algebraically independent elements over $\boldsymbol{k}$. If y_1 is an element of $\boldsymbol{k}[x_1, \cdots, x_n]$ which is not in $\boldsymbol{k}$, then there exist elements $y_2, \cdots, y_n$ of $\boldsymbol{k}[x_1, \cdots, x_n]$ such that 1) $y_i = x_i + x_1^{m_i}$ for some natural number m_i $(i = 2, \cdots, n)$ and 2) $\boldsymbol{k}[x_1, \cdots, x_n]$ is integral over $\boldsymbol{k}[y_1, \cdots, y_n]$ (and therefore $y_1, \cdots, y_n$ are algebraically independent over $\boldsymbol{k}$).*

Remark 1. If $\boldsymbol{k}$ contains infinitely many elements, we may replace the first condition by: "$y_2, \cdots, y_n$ are linear combinations of $x_1, \cdots, x_n$ with coefficients in $\boldsymbol{k}$."

Remark 2. For any given natural number r, the m_i's may be selected so as to be multiple of r, as will easily be seen from the proof below.

Proof. We write y_1 as $\sum_i a_i M_i$, where $a_i \varepsilon \boldsymbol{k}$, $a_i \neq 0$ and the M_i's are monomials in $x_1, \cdots, x_n$. We define weights $m_1 = 1$, $m_2, \cdots, m_n$ of $x_1, x_2, \cdots, x_n$ such that one M_i, say M_1, has greater weight than the others. (For example, set $m_i = (d+1)^{i-1}$ for each i, where d is the degree of the polynomial y_1.) Set $y_i = x_i + x_1^{m_i}$ for $i = 2, \cdots, n$. Then y_1 can be written $a_1 x_1^w + f_1 x_1^{w-1} + \cdots + f_w$, where the f_i's are polynomials in $y_2, \cdots, y_n$ with coefficients in $\boldsymbol{k}$ and $w =$ weight M_1. Then these y_i's are the required elements.

PROPOSITION 1. **(Normalization theorem for polynomial rings)** *Let $\boldsymbol{k}$ be a field and let $x_1, \cdots, x_n$ be algebraically independent elements over $\boldsymbol{k}$. If $\mathfrak{a}$ is an ideal of* rank *r in $\boldsymbol{k}[x_1, \cdots, x_n]$, then there exist elements $y_1, \cdots, y_n$ of $\boldsymbol{k}[x_1, \cdots, x_n]$ such that 1) $\boldsymbol{k}[x_1, \cdots, x_n]$ is integral over $\boldsymbol{k}[y_1, \cdots, y_n]$, 2) $\boldsymbol{k}[y_1, \cdots, y_n] \cap \mathfrak{a}$ is generated by $y_1, \cdots, y_r$ and 3) $y_{r+j} = x_{r+j} + f_j$ with f_j in $\pi[x_1, \cdots, x_r]$ for each $j = 1, \cdots, n-r$, where π is the prime integral domain of $\boldsymbol{k}$.*

Remark 1. This condition 3) shows in particular that $\boldsymbol{k}[x_1, \cdots, x_n] = \boldsymbol{k}[x_1, \cdots, x_r, y_{r+1}, \cdots, y_n]$ and that y_{r+j} $(j \geqq 1)$ is in $\pi[x_1, \cdots, x_n]$.

Remark 2. Assume that $\boldsymbol{k}$ is of characteristic $p \neq 0$. Then taking the natural number r in Remark 2 after Lemma 1 to be equal to p, we see that we can select f_j's to be in $\pi[x_1^p, \cdots, x_r^p]$.

Proof. When $r=0$, our assertion is evident and we prove our assertion by induction on r. Let $\mathfrak{a}'$ be an ideal contained in $\mathfrak{a}$ and of rank $r-1$. Then there exist elements $y_1, \cdots, y_{r-1}, y'_r, \cdots, y'_n$ of $\boldsymbol{k}[x_1, \cdots, x_n]$ which satisfy the conditions in our assertion for $\mathfrak{a}'$ instead of $\mathfrak{a}$. Since $\mathfrak{a}$ is of rank r, $\mathfrak{a} \cap \boldsymbol{k}[y_1, \cdots, y_{r-1}, y'_r, \cdots, y'_n]$ is of rank r ([10, § 8]). Therefore there exists an element y_r of $\mathfrak{a} \cap \boldsymbol{k}[y'_r, \cdots, y'_n]$ which is not zero. Then applying Lemma 1 to y_r and $\boldsymbol{k}[y'_r, \cdots, y'_n]$ we see the existence of $y_{r+1}, \cdots, y_n$ of $\boldsymbol{k}[y'_r, \cdots, y'_n]$ such that i) $y_{r+j} = y'_{r+j} + y'^{m_j}_r$ for some m_j and ii) $\boldsymbol{k}[y'_r, \cdots, y'_n]$ is integral over $\boldsymbol{k}[y_r, \cdots, y_n]$. Then by condition i), we see that condition 3) in our assertion is satisfied by $y_{r+1}, \cdots, y_n$. By condition ii), $\boldsymbol{k}[x_1, \cdots, x_n]$ is integral over $\boldsymbol{k}[y_1, \cdots, y_n]$. Since $\mathfrak{a} \cap \boldsymbol{k}[y_1, \cdots, y_n]$ is of rank r and since $\mathfrak{a}$ contains $y_1, \cdots, y_r$, we see that $\mathfrak{a} \cap \boldsymbol{k}[y_1, \cdots, y_n]$ is generated by $y_1, \cdots, y_r$.

COROLLARY 1. *Let I be an integral domain and let $x_1, \cdots, x_n$ be algebraically independent elements over I. Let $\boldsymbol{k}$ be the field of quotients of I. If $\mathfrak{a}$ is an ideal of $I[x_1, \cdots, x_n]$ such that $\mathfrak{a} \cap I = 0$, then there exist elements $y_1, \cdots, y_n$ of $I[x_1, \cdots, x_n]$ and an element a $(\neq 0)$ of I such that* 1) *$I[a^{-1}, x_1, \cdots, x_n]$ is integral over $I[a^{-1}, y_1, \cdots, y_n]$ and* 2) *$\mathfrak{a}I[a^{-1}, x_1, \cdots, x_n] \cap I[a^{-1}, y_1, \cdots, y_n]$ is generated by $y_1, \cdots, y_r$ with $r = \operatorname{rank} \mathfrak{a}\boldsymbol{k}[x_1, \cdots, x_n]$. Further* 3) *if π is the prime integral domain of I, we can choose $y_{r+1}, \cdots, y_n$ from $\pi[x_1, \cdots, x_n]$.*

Proof. Set $\mathfrak{a}' = \mathfrak{a}\boldsymbol{k}[x_1, \cdots, x_n]$. Take elements $y_1, \cdots, y_n$ as in the above proposition applied to $\mathfrak{a}'$ and $\boldsymbol{k}[x_1, \cdots, x_n]$. Then $y_{r+1}, \cdots, y_n$ are in $\pi[x_1, \cdots, x_n]$ $(\subseteq I[x_1, \cdots, x_n])$. For each $i \leqq r$, there exists an element a_i $(\neq 0)$ of I such that $a_i y_i \in \mathfrak{a}$, because $y_i \in \mathfrak{a}'$. Since $\boldsymbol{k}$ is a field, $a_1 y_1, \cdots, a_r y_r$ are as good as $y_1, \cdots, y_r$. Therefore we may assume that $y_1, \cdots, y_r$ are in $\mathfrak{a}$. Since x_i is integral over $\boldsymbol{k}[y_1, \cdots, y_n]$, there exists an element c_i $(\neq 0)$ of I such that $c_i x_i$ is integral over $I[y_1, \cdots, y_n]$ for each i. On the other hand, let $\mathfrak{p}_1, \cdots, \mathfrak{p}_m$ be the set of prime divisors of $\mathfrak{a}$ which contain nonzero elements of I and let b be a nonzero element of I which is contained in all of $\mathfrak{p}_i$. Let a be the product of b and all the c_j's. Then as is easily seen, this a and the above $y_1, \cdots, y_n$ are the required elements.

COROLLARY 2. **(Normalization theorem (for finitely generated rings))**

Assume that a ring $\mathfrak{o}$ *is generated by elements* $x_1,\cdots,x_n$ *over an integral domain* I. *Assume further that no element* a $(\neq 0)$ *of* I *is a zero-divisor in* $\mathfrak{o}$. *Then there exist elements* $z_1,\cdots,z_t$ *of* $\pi[x_1,\cdots,x_n]$ (*where* π *is the prime integral domain*) *which are algebraically independent over* I *and an element* a $(\neq 0)$ *of* I *such that* $\mathfrak{o}[1/a]$ *is integral over* $I[1/a, z_1,\cdots,z_t]$.

Remark. Observe that if I is a field, then our assumption on I is satisfied and we may take $a=1$.

Proof. Since $\mathfrak{o}$ is a homomorphic image of a polynomial ring $\mathfrak{o}'$ over I, applying Corollary 1 to $\mathfrak{o}'$ and the kernel of the homomorphism, we prove our assertion.

COROLLARY 3. *Let* $\mathfrak{o}$ *be an integral domain which is finitely generated over a field* $\boldsymbol{k}$. *Then for any prime ideal* $\mathfrak{p}$ *of* $\mathfrak{o}$, rank $\mathfrak{p}$ + co-rank $\mathfrak{p}$ *is equal to the transcendence degree of* $\mathfrak{o}$ *over* $\boldsymbol{k}$ *and* co-rank $\mathfrak{p}$ *is equal to the transcendence degree of* $\mathfrak{o}/\mathfrak{p}$ *over* $\boldsymbol{k}$.

Proof. By Corollary 2 and some results in [10, §§ 4-5], we may assume that $\mathfrak{o}$ (and for the last assertion, $\mathfrak{o}/\mathfrak{p}$) is generated by algebraically independent elements of $\boldsymbol{k}$. Then our proposition shows the validity of Corollary 3.

COROLLARY 4. *Let* $\boldsymbol{k}$ *be a field and let* $x_1,\cdots,x_n$ *be elements of a ring containing* $\boldsymbol{k}$. *Then every maximal ideal* $\mathfrak{m}$ *of* $\boldsymbol{k}[x_1,\cdots,x_n]$ *is generated by n elements. Further* $\boldsymbol{k}[x_1,\cdots,x_n]/\mathfrak{m}$ *is algebraic over* $\boldsymbol{k}$. (Zariski [14])

Proof. By Corollary 3, we see that $\boldsymbol{k}[x_1,\cdots,x_n]/\mathfrak{m}$ is algebraic over $\boldsymbol{k}$. Let x'_i be the residue class of x_i modulo $\mathfrak{m}$ for each i and let $f'_i(X_i)$ be the irreducible monic polynomial over $\boldsymbol{k}[x'_1,\cdots,x'_{i-1}]$ which has x'_i as a root. Let f_i be the monic polynomial in x_i with coefficients in $\boldsymbol{k}[x_1,\cdots,x_{i-1}]$ which is obtained from f'_i replacing $x'_1,\cdots,x'_{i-1}, X_i$ by $x_1,\cdots,x_{i-1},x_i$ respectively. Then we see that $\mathfrak{m}$ is generated by $f_1,\cdots,f_n$.

COROLLARY 5. *Let* I *be a field or a Dedekind domain and let* $x_1,\cdots,x_n$ *be algebraically independent elements over* I. *If* $\mathfrak{p}$ *is a prime ideal of rank* r *in the ring* $\mathfrak{o}=I[x_1,\cdots,x_n]$, *then* $\mathfrak{o}_{\mathfrak{p}}$ *is a regular local ring of rank* r.

Proof. We first assume that I is a field. Let $y_1,\cdots,y_n$ be elements of $\mathfrak{o}$ as in Proposition 1 applied to $\mathfrak{p}$. Then $\mathfrak{o}=I[x_1,\cdots,x_r,y_{r+1},\cdots,y_n]$ by condition 3) of the proposition. Let K be the field of quotients of $I[y_{r+1},\cdots,y_n]$. Then $\mathfrak{p}K[x_1,\cdots,x_r]$ is a maximal ideal of $K[x_1,\cdots,x_r]$. Therefore $\mathfrak{p}K[x_1,\cdots,x_r]$ is generated by r elements. Since

$$\mathfrak{o}_{\mathfrak{p}}=K[x_1,\cdots,x_r]_{\mathfrak{p}K[x_1,\cdots,x_r]},$$

this proves our assertion in this case. Now we prove the general case. Set $\mathfrak{q}=\mathfrak{p}\cap I$. If $\mathfrak{q}=0$, then $\mathfrak{o}_{\mathfrak{p}}$ contains the field of quotients of I and the assertion follows from the case where I is a field. Therefore we assume that $\mathfrak{q}\neq 0$. Since $I_{\mathfrak{q}}$ is a principal ideal ring ([10, § 9]) and since $\mathfrak{o}_{\mathfrak{p}}$ contains $I_{\mathfrak{q}}$, we may assume that $\mathfrak{q}$ is generated by an element q. Since $I/\mathfrak{q}$ is a field and since $\mathfrak{p}/q\mathfrak{o}$ is of rank $r-1$, $\mathfrak{p}\mathfrak{o}_{\mathfrak{p}}/q\mathfrak{o}_{\mathfrak{p}}$ is generated by $r-1$ elements and therefore $\mathfrak{p}\mathfrak{o}_{\mathfrak{p}}$ is generated by r elements, which proves our assertion.

COROLLARY 6. *Let I be an integral domain and let $x_1,\cdots,x_n$ $(n\geqq 1)$ be algebraically independent elements over I. Then there exists a maximal ideal $\mathfrak{m}$ of $I[x_1,\cdots,x_n]$ such that $\mathfrak{m}\cap I=0$ if and only if there exists an element a $(\neq 0)$ of I such that $I[1/a]$ is a field.*

Proof. If there exists such an a, then a maximal ideal $\mathfrak{m}$ containing ax_1-1 meets I only in 0. Conversely, we assume that there exists such a maximal ideal $\mathfrak{m}$. Set $\mathfrak{o}=I[x_1,\cdots,x_n]/\mathfrak{m}$. Then by Corollary 2, there exists an element a $(\neq 0)$ of I such that for a suitable system of algebraically independent elements $y_1,\cdots,y_r$ of $\mathfrak{o}$ over I, $\mathfrak{o}[1/a]$ is integral over $I[1/a,y_1,\cdots,y_r]$. Since $\mathfrak{o}$ is a field, $\mathfrak{o}[1/a]=\mathfrak{o}$. Further, since there is a field which is integral over $I[1/a,y_1,\cdots,y_r]$, $I[1/a,y_1,\cdots,y_r]$ must be a field ([10, § 4]), whence $r=0$ and $I[1/a]$ is a field (and therefore this is the field of quotients of I).

We have proved at the same time the following

COROLLARY 7. *Assume that $\mathfrak{o}$ is a finitely generated ring over a ring I. If $\mathfrak{m}$ is a maximal ideal of $\mathfrak{o}$, then $\mathfrak{o}/\mathfrak{m}$ is algebraic over $I/(\mathfrak{m}\cap I)$.*

PROPOSITION 2. *Let $x_1,\cdots,x_n$ be algebraically independent elements over a Noetherian integral domain I. If I satisfies the finiteness condition for integral extensions, then so does $I[x_1,\cdots,x_n]$.*

Proof. Let L be a finite extension field of the field of quotients K of $I[x_1,\cdots,x_n]$; we have only to show that every integral extension of $I[x_1,\cdots,x_n]$ contained in L is a finite $I[x_1,\cdots,x_n]$-module. By our assumption on I, we may assume that I is normal. If L is separable over K, then the assertion is obvious ([10, § 5]). When L is inseparable over K, take elements $a_1,\cdots,a_r$ of I and a power q of the characteristic of I such that

$$L'=L(a_1{}^{1/q},\cdots,a_r{}^{1/q},x_1{}^{1/q},\cdots,x_n{}^{1/q})$$

is separable over

$$K'=K(a_1{}^{1/q},\cdots,a_r{}^{1/q},x_1{}^{1/q},\cdots,x_n{}^{1/q}).$$

Let I' be the derived normal ring of $I[a_1^{1/q}, \cdots, a_r^{1/q}]$. Then I' is finite over I by our assumption, and $I'[x_1^{1/q}, \cdots, x_n^{1/q}]$ is finite over $I[x_1, \cdots, x_n]$. Since L' is separable over K' the integral closure of $I'[x_1^{1/q}, \cdots, x_n^{1/q}]$ in L' is finite over $I'[x_1^{1/q}, \cdots, x_n^{1/q}]$, and is therefore finite over $I[x_1, \cdots, x_n]$. Since $I[x_1, \cdots, x_n]$ is Noetherian, we see that every integral extension of $I[x_1, \cdots, x_n]$ contained in L is finite and we have proved the assertion.

COROLLARY. *If an integral domain $\mathfrak{o}$ is finitely generated over a field, then the derived normal ring of $\mathfrak{o}$ is a finite $\mathfrak{o}$-module.*

2. Definition of spots. An integral domain $\mathfrak{o}$ is called an *affine ring* over a *ground ring* I if I is a field or a Dedekind domain and if $\mathfrak{o}$ is finitely generated over I; here we assume that any ground ring satisfies the finiteness condition for integral extensions (cf. the remark "On the restriction of ground rings" at the beginning of this paper).

Remark. If $\mathfrak{p}$ is a prime ideal of an affine ring $\mathfrak{o}$ over a ground ring I, then $\mathfrak{o}/\mathfrak{p}$ is an affine ring over $I/(\mathfrak{p} \cap I)$; this last ring is I itself or a field and satisfies the finiteness condition for integral extensions.

A field L is called a *function field* over a ground ring I if there exists an affine ring $\mathfrak{o}$ over I such that L is the field of quotients of $\mathfrak{o}$; such $\mathfrak{o}$ is called an affine ring of L.

A ring P is called a *spot* over a ground ring I if there exists an affine ring $\mathfrak{o}$ over I which has a prime ideal $\mathfrak{p}$ such that $P = \mathfrak{o}_{\mathfrak{p}}$; if L is the field of quotients of P, P is called a spot of L.

Remark 1. If P is a spot over a ground ring I and if $\mathfrak{m}$ is the maximal ideal of P, then P is a spot over $I_{(\mathfrak{m} \cap I)}$; this last ring is a field or a discrete valuation ring ([10, § 9]).

Remark 2. If $\mathfrak{p}$ is a prime ideal of a spot P over a ground ring I, then 1) $P_{\mathfrak{p}}$ is a spot over I and 2) $P/\mathfrak{p}$ is a spot over $I/(\mathfrak{p} \cap I)$.

Remark 3. A spot P is a (Noetherian) local integral domain.

A subring B of a spot P is called a *basic ring* of P if 1) P is a spot over B, 2) B is a field or a valuation ring and P dominates B and 3) the residue class field of P is a finite algebraic extension of that of B.

A basic ring which is a field is called a *basic field.*

PROPOSITION 3. *Every spot P has a basic ring.*

Proof. Let I be a ground ring of P; by Remark 1, we may assume that

I is a field or a discrete valuation ring dominated by P. Let p be either a prime element of I or zero according to whether I is a valuation ring or a field. Let $\mathfrak{p}$ be the maximal ideal of P and let $x_1, \cdots, x_d$ be elements of P whose residue classes modulo $\mathfrak{p}$ form a transcendence base of $P/\mathfrak{p}$ over I/pI. Since I is a field or a valuation ring, $x_1, \cdots, x_d$ are algebraically independent over I, $pI[x_1, \cdots, x_d]$ is a prime ideal and $\mathfrak{p} \cap I[x_1, \cdots, x_d] = pI[x_1, \cdots, x_d]$. Set $B = I(x_1, \cdots, x_d)$. Then B is a field or a valuation ring which is dominated by P and the residue class field of P is a finite algebraic extension of that of B. Further B satisfies the finiteness condition for integral extensions by virtue of Proposition 2. Therefore B is a basic ring of P.

A spot P is said to be of *the first kind* if it has a basic field; otherwise, P is said to be of *the second kind*.

Remark. A spot of the first kind may have a basic ring which is not a field.

3. Dimension and rank of spots. Let L be a function field over a ground ring I. The *dimension* of L over I (in symbols, $\dim_I L$ or merely $\dim L$) is n or $n+1$ according to whether I is a field or not, where n is the transcendence degree of L over I. The *dimension* of a spot P over a ground ring I (in symbols, $\dim_I P$ or merely $\dim P$) is defined to be the dimension of the residue class field of P modulo its maximal ideal $\mathfrak{m}$ over $I/(\mathfrak{m} \cap I)$.

LEMMA 1. *Let $x_1, \cdots, x_n$ be algebraically independent elements over a Dedekind domain I and let $\mathfrak{m}$ be a maximal ideal of the ring $\mathfrak{o} = I[x_1, \cdots, x_n]$. Assume that $\mathfrak{m} \cap I \neq 0$. Then $\mathfrak{o}_\mathfrak{m}$ is a regular local ring of rank $n+1$; if $\mathfrak{q}$ is a prime ideal contained in $\mathfrak{m}$, then* rank $\mathfrak{q}$ + co-rank $\mathfrak{q} = n+1$ *and* co-rank $\mathfrak{q}$ = co-rank $\mathfrak{q}\mathfrak{o}_\mathfrak{m}$.

Proof. The first assertion follows from Corollary 5 to Proposition 1. Since $\mathfrak{o}_\mathfrak{m}$ is a regular local ring of rank $n+1$, rank $\mathfrak{q}\mathfrak{o}_\mathfrak{m}$ + co-rank $\mathfrak{q}\mathfrak{o}_\mathfrak{m} = n+1$ (by a result due to Krull [5]; a proof will be given in the appendix at the end of the present paper). By the definition of rank, rank $\mathfrak{q}$ = rank $\mathfrak{q}\mathfrak{o}_\mathfrak{m}$. Since every maximal ideal of $\mathfrak{o}$ is at most of rank $n+1$, co-rank $\mathfrak{q} \leqq n+1$ — rank $\mathfrak{q}$, while, by the definition of co-rank, co-rank $\mathfrak{q} \geqq$ co-rank $\mathfrak{q}\mathfrak{o}_\mathfrak{m}$. Therefore co-rank $\mathfrak{q}$ = co-rank $\mathfrak{q}\mathfrak{o}_\mathfrak{m}$ and rank $\mathfrak{q}$ + co-rank $\mathfrak{q} = n+1$.

THEOREM 1. *Let $\mathfrak{o}$ be an affine ring over a ground ring I. Let $0 = \mathfrak{p}_0 \subset \mathfrak{p}_1 \subset \cdots \subset \mathfrak{p}_r$ be a maximal chain of prime ideals $\mathfrak{p}_i$ (that is, each $\mathfrak{p}_i/\mathfrak{p}_{i-1}$ is of rank 1 and $\mathfrak{p}_r$ is maximal). If $\mathfrak{p}_r \cap I = 0$, then r is equal to the*

transcendence degree of $\mathfrak{o}$ *over* I; *if* $\mathfrak{p}_r \cap I \neq 0$, *then* $r-1$ *is equal to the transcendence degree of* $\mathfrak{o}$ *over* I.

Proof. When $\mathfrak{p}_r \cap I = 0$, we may assume that I is a field. Then our assertion follows easily from Corollary 3 to Proposition 1. Therefore we assume that $\mathfrak{p}_r \cap I \neq 0$. Let $x_1, \cdots, x_n$ be algebraically independent elements over I such that there exists a homomorphism ϕ from $I[x_1, \cdots, x_n]$ onto $\mathfrak{o}$; let $\mathfrak{q}$ be the kernel of ϕ. Then we have rank $\mathfrak{q}$ + rank $\mathfrak{o} = n+1$, by Lemma 1. Let $\boldsymbol{k}$ be the field of quotients of I. Since $\mathfrak{q} \cap I = 0$, $I[x_1, \cdots, x_n]_{\mathfrak{q}}$ contains $\boldsymbol{k}$. Therefore we have $n - \operatorname{rank} \mathfrak{q}$ = transcendence degree of $\mathfrak{o}$ over I; we denote this number by t. Since $\operatorname{rank} \mathfrak{q} + \operatorname{rank} \mathfrak{o} = n+1$, we have $\operatorname{rank} \mathfrak{o} = t+1$. Therefore $r \leqq t+1$. Therefore, when $t=0$, our assertion is evident. Thus we will prove our assertion by induction on t. Assume that $t \geqq 1$. By Lemma 1, $\operatorname{rank} \mathfrak{o} = \operatorname{rank} \mathfrak{o}_{\mathfrak{p}_r}$. Therefore $r > 1$.

1) When $\mathfrak{p}_{r-1} \cap I \neq 0$: $\mathfrak{o}/\mathfrak{p}_{r-1}$ is an affine ring over the field $I/(\mathfrak{p}_{r-1} \cap I)$, which shows that $\mathfrak{o}/\mathfrak{p}_{r-1}$ has transcendence degree 1 over $I/(\mathfrak{p}_{r-1} \cap I)$. Therefore there exist a basic ring B (defined even in the non-restricted case) of $\mathfrak{o}_{\mathfrak{p}_{r-1}}$ which is of transcendence degree 1 over I (as in the proof of Proposition 3). Set $\mathfrak{o}' = B[\mathfrak{o}]$. Then $0 = \mathfrak{p}_0\mathfrak{o}' \subset \mathfrak{p}_1\mathfrak{o}' \subset \cdots \subset \mathfrak{p}_{r-1}\mathfrak{o}'$ is a maximal chain of prime ideals of $\mathfrak{o}'$, because $\mathfrak{o}'$ is a ring of quotients of $\mathfrak{o}$ (see the construction of B in the proof of Proposition 3); that $\mathfrak{p}_{r-1}\mathfrak{o}'$ is maximal follows from the fact that $\mathfrak{o}/\mathfrak{p}_{r-1}$ is an affine ring over the field $I/(\mathfrak{p}_{r-1} \cap I)$. Since $\mathfrak{o}'$ has transcendence degree $t-1$ over B, we have $r-1=t$ by our induction assumption. Thus, this case is settled.

2) When $\mathfrak{p}_{r-1} \cap I = 0$: Since $r > 1$, $\mathfrak{o}/\mathfrak{p}_{r-1}$ has a transcendence degree less than that of $\mathfrak{o}$ over I. Therefore by our induction assumption, $\mathfrak{o}/\mathfrak{p}_{r-1}$ is algebraic over I. Set $\mathfrak{o}' = \boldsymbol{k}[\mathfrak{o}]$. Then $0 = \mathfrak{p}_0\mathfrak{o}' \subset \mathfrak{p}_1\mathfrak{o}' \subset \cdots \subset \mathfrak{p}_{r-1}\mathfrak{o}'$ is a maximal chain of prime ideal in $\mathfrak{o}'$. Therefore $r-1=t$ by Corollary 3 to Proposition 1. Thus we have settled this case, too.

Corollary 1. *Let P be a spot of a function field L. If I is a basic ring of P, then* $\dim_I L = \operatorname{rank} P$.

Corollary 2. *If $\mathfrak{p}$ is a prime ideal of a spot P, then* $\operatorname{rank} \mathfrak{p} + \text{co-rank}\, \mathfrak{p} = \operatorname{rank} P$.

Corollary 3. *Let $\mathfrak{o}$ be an affine ring of a function field over a ground ring I. If $\mathfrak{p}$ is a prime ideal of $\mathfrak{o}$, then* $\operatorname{rank} \mathfrak{p} + \dim_I \mathfrak{o}_{\mathfrak{p}} = \dim_I L$.

4. Analytical unramifiedness of spots. A semi-local integral domain is said to be *analytically unramified* if its completion has no niloptent elements.

A prime ideal $\mathfrak{p}$ of a semi-local ring $\mathfrak{o}$ is said to be *analytically unramified* if $\mathfrak{o}/\mathfrak{p}$ is analytically unramified.

LEMMA 1. *Let $\mathfrak{o}$ be a normal semi-local ring. Assume that a prime ideal $\mathfrak{p}$ of rank 1 in $\mathfrak{o}$ is analytically unramified. Let $\mathfrak{o}^*$ be the completion of $\mathfrak{o}$. Then for every prime divisor $\mathfrak{p}^*$ of $\mathfrak{p}\mathfrak{o}^*$, $\mathfrak{o}^*_{\mathfrak{p}^*}$ is a valuation ring.* (Zariski [15])

Remark. This result does not show that $\mathfrak{o}^*$ is an integral domain, but shows that $\mathfrak{p}^*$ contains only one prime divisor $\mathfrak{P}^*$ of zero in $\mathfrak{o}^*$ and that the primary component of zero belonging to $\mathfrak{P}^*$ coincides with $\mathfrak{P}^*$ (and the ring of quotients of $\mathfrak{o}^*/\mathfrak{P}^*$ with respect to $\mathfrak{p}^*/\mathfrak{P}^*$ is a valuation ring).

Proof. Let w be an element of $\mathfrak{p}$ which is not in $\mathfrak{p}^2\mathfrak{o}_{\mathfrak{p}}$. Since $\mathfrak{o}$ is a normal ring, $\mathfrak{o}_{\mathfrak{p}}$ is a discrete valuation ring ([10, § 9]) and $w\mathfrak{o}:\mathfrak{p}$ is not contained in $\mathfrak{p}$. Let b be an element of $w\mathfrak{o}:\mathfrak{p}$ which is not in $\mathfrak{p}$ and let a^* be an element of $\mathfrak{o}^*$ which is not in $\mathfrak{p}^*$ but is in every other prime divisor of $\mathfrak{p}\mathfrak{o}^*$ (such a^* exists because $\mathfrak{p}$ is analytically unramified). Set $c^*=a^*b$. Then c^* is not in $\mathfrak{p}^*$ and $\mathfrak{p}^*c^*\subseteq\mathfrak{o}^*\mathfrak{p}b\subseteq w\mathfrak{o}^*$. Therefore we see that $\mathfrak{p}^*\mathfrak{o}^*_{\mathfrak{p}^*}\subseteq w\mathfrak{o}^*_{\mathfrak{p}^*}$, whence $\mathfrak{p}^*\mathfrak{o}^*_{\mathfrak{p}^*}=w\mathfrak{o}^*_{\mathfrak{p}^*}$. Since w is not a zero-divisor in $\mathfrak{o}^*$ (Corollary to Lemma 0.4), $\mathfrak{p}^*$ properly contains a prime divisor $\mathfrak{P}^*$ of zero. Since in any Noetherian ring, a principal ideal generated by an element of its J-radical cannot properly contain any prime ideal other than zero ([10, § 6]), we see that $\mathfrak{P}^*$ must be the kernel of the natural homomorphism from $\mathfrak{o}^*$ into $\mathfrak{o}^*_{\mathfrak{p}^*}$. Now, since $\mathfrak{p}^*\mathfrak{o}^*_{\mathfrak{p}^*}$ is principal and since $\mathfrak{o}^*_{\mathfrak{p}^*}$ is a local integral domain which is not a field, $\mathfrak{o}^*_{\mathfrak{p}^*}$ must be a valuation ring.

LEMMA 2. *Let $\mathfrak{o}$ be a normal semi-local ring and let $\mathfrak{o}^*$ be its completion. Assume that for an element t of $\mathfrak{o}$, which is neither zero nor unit in $\mathfrak{o}$, every prime divisor of $t\mathfrak{o}$ is analytically unramified. If an element u is integral over $\mathfrak{o}^*$ and if tu is in $\mathfrak{o}^*$, then u is also in $\mathfrak{o}^*$.* (Zariski [16])

Proof. Let $\mathfrak{p}_1,\cdots,\mathfrak{p}_r$ be all the prime divisors of $t\mathfrak{o}$; they are of rank 1 ([10, § 9]). Let $\mathfrak{p}^*_{ij}$ $(j=1,\cdots,n(i))$ be all the prime divisors of $\mathfrak{p}_i\mathfrak{o}^*$. We set $S=\{a;a\in\mathfrak{o},a\not\in\mathfrak{p}_i$ for every $i\}$.[2] Then since $\mathfrak{o}$ is a normal ring, $\mathfrak{o}_S$ is a semi-local Dedekind domain and is a principal ideal ring ([10, § 9]). Let x_i be an element of $\mathfrak{o}$ such that $\mathfrak{p}_i\mathfrak{o}_S=x_i\mathfrak{o}_S$. Let e_i be natural numbers such that $t\mathfrak{o}_S=x_1^{e_1}\cdots x_r^{e_r}\mathfrak{o}_S$. Now we have only to show that for some element s of S, tus is divisible by $x_1^{e_1}\cdots x_r^{e_r}$ in $\mathfrak{o}^*$. For, once this is done, the proof concludes as follows: since $t\mathfrak{o}:s\mathfrak{o}=t\mathfrak{o}$, we see that $t\mathfrak{o}^*:s\mathfrak{o}^*=t\mathfrak{o}^*$ (Lemma 0.4). Since $x_1^{e_1}\cdots x_r^{e_r}\mathfrak{o}\subseteq t\mathfrak{o}$, we see that $x_1^{e_1}\cdots x_r^{e_r}\mathfrak{o}^*\subseteq t\mathfrak{o}^*$.

[2] The notation $\{a;P\}$ denotes the set of a which satisfy P.

Therefore the fact that $tus \varepsilon x_1^{e_1} \cdots x_r^{e_r}\mathfrak{o}^*$ shows that $tus \varepsilon t\mathfrak{o}^*$ and tu is in $t\mathfrak{o}^* : s\mathfrak{o}^* = t\mathfrak{o}^*$. Then since t is not a zero-divisor in $\mathfrak{o}^*$ (Corollary to Lemma 0.4), u is in $\mathfrak{o}^*$. Now we proceed to show that tus is in $x_1^{e_1} \cdots x_r^{e_r}\mathfrak{o}^*$ for some s in S. Let w_{ij} be the valuation of the field of quotients of $\mathfrak{o}^*_{\mathfrak{p}^*_{ij}}$ with $\mathfrak{o}^*_{\mathfrak{p}^*_{ij}}$ as valuation ring and $w_{ij}(x_i) = 1$. Let ϕ_{ij} be the natural homomorphism from $\mathfrak{o}^*$ into $\mathfrak{o}^*_{\mathfrak{p}^*_{ij}}$. Let $f_1, \cdots, f_r$ be non-negative integers satisfying the following condition: tus is in $x_1^{f_1} \cdots x_r^{f_r}\mathfrak{o}^*$ for some s of S but for any s of S and for any i, tus is not in $x_1^{f_1} \cdots x_{i-1}^{f_{i-1}} x_i^{f_i+1} x_{i+1}^{f_{i+1}} x_r^{f_r}\mathfrak{o}^*$. Then it is sufficient to show that $f_i \geqq e_i$. Assume the contrary, for instance that $f_1 < e_1$. We take an element s of S such that tus is in $x_1^{f_1} \cdots x_r^{f_r}\mathfrak{o}^*$ and take an element z of $\mathfrak{o}^*$ such that $tus = x_1^{f_1} \cdots x_r^{f_r} z$. Since the kernel of ϕ_{ij} is a prime ideal and since $\mathfrak{o}^*_{\mathfrak{p}^*_{ij}}$ is a normal ring, we may regard ϕ_{ij} as a homomorphism of the total quotient ring $\mathfrak{o}^{*\prime}$ of $\mathfrak{o}^*$ into the field of quotients of $\mathfrak{o}^*_{\mathfrak{p}^*_{ij}}$. Then since u is integral over $\mathfrak{o}^*$, $\phi_{ij}(u)$ is in $\mathfrak{o}^*_{\mathfrak{p}^*_{ij}}$, whence $w_{ij}(\phi_{ij}(u)) \geqq 0$. Since $w_{1j}(\phi_{1j}(t)) = e_1 > f_1 = w_{1j}(\phi_{1j}(x_1^{f_1} \cdots x_r^{f_r}))$, we have $w_{1j}(\phi_{1j}(z)) \geqq 1$. This shows that $\phi_{1j}(z)$ is in $\phi_{1j}(\mathfrak{p}^*_{1j})$ and therefore z is in $\mathfrak{p}^*_{1j}$ for every j because z is in $\mathfrak{o}^*$ (observe that the kernel of ϕ_{1j} is contained in $\mathfrak{p}^*_{1j}$). Since $x_1\mathfrak{o}^*_S = \bigcap_j \mathfrak{p}^*_{1j}\mathfrak{o}^*_S$, z is in $x_1\mathfrak{o}^*_S$ and therefore there exists an element s' of S such that $zs' = x_1 z'$ with an element z' of $\mathfrak{o}^*$. Thus we see that $tus'' = x_1^{f_1+1} x_2^{f_2} \cdots x_r^{f_r} z'$ with $s'' = ss'$ (which is in S). This is a contradiction and we have $e_i \leqq f_i$ for every i. Thus the lemma is proved.

LEMMA 3. *Let $\mathfrak{r}$ be a normal local ring whose completion is a normal ring. Let L be a finite separable extension of the field of quotients R of $\mathfrak{r}$ and let $\mathfrak{J}$ be the integral closure of $\mathfrak{r}$ in L. Assume that every prime ideal $\mathfrak{p}$ of rank 1 in $\mathfrak{J}$ is analytically unramified. Then the completion of $\mathfrak{J}$ is integrally closed, that is, for every maximal ideal $\mathfrak{m}$ of $\mathfrak{J}$, the completion of $\mathfrak{J}_\mathfrak{m}$ is a normal ring.* (Zariski [16])

Proof. Let a be an element of $\mathfrak{J}$ such that $L = R(a)$. Let d be the discriminant of the irreducible monic polynomial over $\mathfrak{r}$ which has a as a root. On the other hand, let $\mathfrak{r}^*$ and $\mathfrak{J}^*$ be the completions of $\mathfrak{r}$ and $\mathfrak{J}$ respectively and denote by $\mathfrak{J}^{*\prime}$ the integral closure of $\mathfrak{J}^*$ in its total quotient ring. Since $\mathfrak{J}$ is a finite $\mathfrak{r}$-module, the integral closure of $\mathfrak{r}^*[a]$ in its total quotient ring coincides with $\mathfrak{J}^{*\prime}$ (Lemma 0.6). Therefore we see that $d\mathfrak{J}^{*\prime} \subseteq \mathfrak{r}^*[a]$ ([10, §5]). By Lemma 2, we see that $\mathfrak{J}^{*\prime} = \mathfrak{J}^*$ and therefore $\mathfrak{J}^*$ is integrally closed. Since $\mathfrak{J}^*$ is integrally closed and is Noetherian, we see that $\mathfrak{J}^*$ is the direct sum of normal rings. On the other hand, since $\mathfrak{J}^*$ is the direct sum of completions of $\mathfrak{J}_\mathfrak{m}$, where $\mathfrak{m}$ runs over all maximal ideals of $\mathfrak{J}$ (Lemma 0.2), we see that the completion of $\mathfrak{J}_\mathfrak{m}$ is a normal ring for every maximal ideal $\mathfrak{m}$ of $\mathfrak{J}$.

LEMMA 4. *Let $\mathfrak{o}$ be a normal local ring and assume that its completion $\mathfrak{o}^*$ is an integral domain. Let L be the field of quotients of $\mathfrak{o}$. Assume that a local ring $\mathfrak{o}'$ which is a subring of L satisfies following conditions:* 1) *$\mathfrak{o}'$ dominates $\mathfrak{o}$,* 2) *$\mathfrak{o}'/\mathfrak{m}'$ is a finite algebraic extension of $\mathfrak{o}/\mathfrak{m}$, where $\mathfrak{m}$ and $\mathfrak{m}'$ denote the maximal ideals of $\mathfrak{o}$ and $\mathfrak{o}'$ respectively,* 3) *$\mathfrak{m}\mathfrak{o}'$ is a primary ideal belonging to $\mathfrak{m}'$ and* 4) $\operatorname{rank}\mathfrak{o}' = \operatorname{rank}\mathfrak{o}$. *Then $\mathfrak{o}$ coincides with $\mathfrak{o}'$.*

Proof. Let $\mathfrak{o}'^*$ be the completion of $\mathfrak{o}'$. Since $\mathfrak{m}' \cap \mathfrak{o} = \mathfrak{m}$, we see that $\mathfrak{m}^i \subseteq \mathfrak{o} \cap \mathfrak{m}'^i$ for every i. Therefore there exists a natural homomorphism ϕ from $\mathfrak{o}^*$ into $\mathfrak{o}'^*$. Set $\mathfrak{o}^{**} = \phi(\mathfrak{o}^*)$, (this is the closure of $\mathfrak{o}$ in $\mathfrak{o}'^*$ in the topology of $\mathfrak{o}'^*$). By conditions 2) and 3), we see that $\mathfrak{o}'^*$ is a finite $\mathfrak{o}^{**}$-module (Lemma 0.5). Therefore $\operatorname{rank}\mathfrak{o}^{**} = \operatorname{rank}\mathfrak{o}'^*$ ([10, § 8]). Since $\operatorname{rank}\mathfrak{o} = \operatorname{rank}\mathfrak{o}'$, $\operatorname{rank}\mathfrak{o}^* = \operatorname{rank}\mathfrak{o}'^*$ (Lemma 0.1) and $\operatorname{rank}\mathfrak{o}^* = \operatorname{rank}\mathfrak{o}^{**}$. Since $\mathfrak{o}^*$ is an integral domain, we see that ϕ is an isomorphism. Thus we see that $\mathfrak{o}$ is a subspace of $\mathfrak{o}'$ and $\mathfrak{o}'^*$ is integral over $\mathfrak{o}^*$. Now let a/b $(a, b \in \mathfrak{o})$ be an element of $\mathfrak{o}'$. Since a/b is integral over $\mathfrak{o}^*$, there exist elements $c^*_1, \cdots, c^*_n$ of $\mathfrak{o}^*$ such that $(a/b)^n + c^*_1(a/b)^{n-1} + \cdots + c^*_n = 0$ and therefore $a^n + ba^{n-1}c^*_1 + \cdots + b^n c^*_n = 0$, which shows that a^n is in the ideal generated by $ba^{n-1}, \cdots, b^n$. Since $(\sum_1^n b^i a^{n-i}\mathfrak{o}^*) \cap \mathfrak{o} = \sum_1^n b^i a^{n-i}\mathfrak{o}$ (Lemma 0.1), we see that there exist elements $c_1, \cdots, c_n$ of $\mathfrak{o}$ such that $a^n + ba^{n-1}c_1 + \cdots + b^n c_n = 0$. Therefore $(a/b)^n + c_1(a/b)^{n-1} + \cdots + c_n = 0$, that is, a/b is integral over $\mathfrak{o}$. Since $\mathfrak{o}$ is normal, a/b is in $\mathfrak{o}$, which proves Lemma 4.

Now we come to the important

THEOREM 2. *Any spot is analytically unramified. Further the completion of a normal spot is a normal ring.*

Proof. Let P be a spot of rank r. We prove our assertion by induction on r. When $r = 0$, our assertion is obvious. Assume that $r > 0$ and that the theorem is true for spots of rank $\leqq r - 1$. We first remark the following fact. Let $\mathfrak{o}$ be a semi-local integral domain. If for every maximal ideal $\mathfrak{m}$ of $\mathfrak{o}$, $\mathfrak{o}_\mathfrak{m}$ is analytically unramified, then $\mathfrak{o}$ is also analytically unramified (the proof follows easily from the fact that the completion of $\mathfrak{o}$ is the direct sum of completions of rings $\mathfrak{o}_\mathfrak{m}$). Therefore our induction assumption means that if $\mathfrak{o}$ is a semi-local integral domain of rank r such that for any maximal ideal $\mathfrak{m}$ of $\mathfrak{o}$, $\mathfrak{o}_\mathfrak{m}$ is a spot and if $\mathfrak{p}$ is a nonzero prime ideal of $\mathfrak{o}$, then $\mathfrak{p}$ is analytically unramified. Let B be a basic ring of P and let $x_1, \cdots, x_r$ be a system of parameters of P, where if B is not a field, we choose x_1 to be a prime element of B. By Corollary 1 to Theorem 1, the field of quotients L

of P is algebraic over $B[x_1, \cdots, x_r]$. Let $\mathfrak{J}$ be the integral closure of $B[x_1, \cdots, x_r]$ in L. Further set $P' = P[\mathfrak{J}]$. $\mathfrak{J}$ is a finite $B[x_1, \cdots, x_r]$-module by Proposition 2 and P' is a finite P-module.

(a) When L is separable over $B[x_1, \cdots, x_r]$, let $\mathfrak{m}'$ be an arbitrary maximal ideal of P' and set $\mathfrak{m} = \mathfrak{m}' \cap \mathfrak{J}$. Then by Lemma 3 and by our induction assumption, we see that the completion of $\mathfrak{J}_{\mathfrak{m}}$ is a normal ring. Since $x_1, \cdots, x_r$ is a system of parameters of $P'_{\mathfrak{m}'}$ and of $\mathfrak{J}_{\mathfrak{m}}$, we see that $P'_{\mathfrak{m}'} = \mathfrak{J}_{\mathfrak{m}}$ by Lemma 4 and Theorem 1. Since the completion of P' is the direct sum of the completions of the rings $P'_{\mathfrak{m}'}$, P' is analytically unramified. Since P' is a finite P-module, P is a subspace of P' (Lemma 0.6) and P is also analytically unramified. When P is normal, P' coincides with P and its completion is a normal ring.

(b) Next we assume that L is not separable over $B[x_1, \cdots, x_r]$. Take elements $a_1, \cdots, a_s$ of B and a power q of the characteristic of B such that

$$L' = L(a_1^{1/q}, \cdots, a_s^{1/q}, x_1^{1/q}, \cdots, x_r^{1/q})$$

is separable over

$$B[a_1^{1/q}, \cdots, a_s^{1/q}, x_1^{1/q}, \cdots, x_r^{1/q}].$$

Let B' be the derived normal ring of $B[a_1^{1/q}, \cdots, a_s^{1/q}]$ and let $\mathfrak{J}'$ be the integral closure of $B[x_1, \cdots, x_r]$ in L'. Since L' is separable over $B'[x^{1/q}, \cdots, x_r^{1/q}]$, for every maximal ideal $\mathfrak{m}''$ of $P'' = P[\mathfrak{J}']$ the completion of $P''_{\mathfrak{m}''}$ is a normal ring and $P''_{\mathfrak{m}''} = \mathfrak{J}'_{(\mathfrak{m}'' \cap \mathfrak{J}')}$ (by our observation in (a) above). By Proposition 2, $\mathfrak{J}'$ is a finite $B[x_1, \cdots, x_r]$-module and P'' is a finite P-module, which shows that P is a subspace of P''. Therefore we see that P is also analytically unramified. Assume that P is normal. Let $b_1, \cdots, b_t$ be elements of $\mathfrak{J}$ which are maximally linearly independent over $B[x_1, \cdots, x_r]$ and such that the module generated by $b_1, \cdots, b_t$ over $B[x_1, \cdots, x_r]$ is a ring; let $c_1, \cdots, c_u$ be elements of $\mathfrak{J}'$ which are maximally linearly independent over $B[x_1, \cdots, x_r, b_1, \cdots, b_t]$ and such that the module generated by $c_1, \cdots, c_u$ over $B[x_1, \cdots, x_r, b_1, \cdots, b_t]$ is a ring. Since $\mathfrak{J}'$ is a finite $B[x_1, \cdots, x_r]$-module, there exists an element d $(\neq 0)$ of $B[x_1, \cdots, x_r]$ such that $d\mathfrak{J}' \subseteq B[x_1, \cdots, x_r, b_1, \cdots, b_t, c_1, \cdots, c_u]$. Let P^*, P''^* and $\mathfrak{r}^*$ be the completions of P, P'' and $B[x_1, \cdots, x_r]_{(x_1, \cdots, x_r)}$ respectively. Then we have $dP''^* \subseteq \mathfrak{r}^*[b_1, \cdots, b_t, c_1, \cdots, c_u]$. Let $\mathfrak{S}$ be the integral closure of P^* in its total quotient ring. Since P''^* is integrally closed, $\mathfrak{S}$ is contained in P''^*. Therefore $d\mathfrak{S} \subseteq \mathfrak{r}^*[b_1, \cdots, b_t, c_1, \cdots, c_u]$. Since $c_1, \cdots, c_u$ are linearly independent over $\mathfrak{r}^*[b_1, \cdots, b_t]$ (Lemma 0.6), we see that $d\mathfrak{S} \subseteq \mathfrak{r}^*[b_1, \cdots, b_t]$ and $d\mathfrak{S} \subseteq P^*$. Therefore by Lemma 2 and by our induction assumption, we have $P^* = \mathfrak{S}$ and P^* is a normal ring. Thus the proof is completed.

COROLLARY. *If P is a spot, then the derived normal ring of P is a finite P-module.*

Proof. The ring P' in the above proof is a normal ring and is a finite module over P. Therefore P' is the derived normal ring of P, which proves our assertion.

Remark. It is known that if a semi-local integral domain $\mathfrak{o}$ is analytically unramified, then the derived normal ring of $\mathfrak{o}$ is a finite $\mathfrak{o}$-module. (Nagata [8]; a proof will also be given in the second paper of the present sequence.)

5. Finiteness of derived normal rings of affine rings.

THEOREM 3. *If $\mathfrak{o}$ is an affine ring, then the derived normal ring $\mathfrak{o}'$ of $\mathfrak{o}$ is a finite $\mathfrak{o}$-module.*

Proof. Let I be a groud ring of $\mathfrak{o}$. Then by Corollary 2 to Proposition 1, there exist an element a $(\neq 0)$ of I and elements $y_1, \cdots, y_n$ of $\mathfrak{o}$ which are algebraically independent over I such that $\mathfrak{o}[1/a]$ is integral over $I[1/a, y_1, \cdots, y_n]$. Let $\mathfrak{o}''$ be the integral closure of $I[1/a, y_1, \cdots, y_n]$ in the field of quotients L of $\mathfrak{o}$. Then by Proposition 2, $\mathfrak{o}''$ is a finite $I[1/a, y_1, \cdots, y_n]$-module. If a is a unit in $\mathfrak{o}$, then $\mathfrak{o}'' = \mathfrak{o}'$ and $\mathfrak{o}'$ is a finite $\mathfrak{o}$-module. Therefore we treat the case where a is non-unit in $\mathfrak{o}$. Since $\mathfrak{o}''$ is a finite $I[1/a, y_1, \cdots, y_n]$-module, there exist a finite number of elements $c_1, \cdots, c_r$ in $\mathfrak{o}'$ such that $\mathfrak{o}'' = \mathfrak{o}[1/a, c_1, \cdots, c_r]$ (because $\mathfrak{o}'' = \mathfrak{o}'[1/a]$). Set $\mathfrak{o}_1 = \mathfrak{o}[c_1, \cdots, c_r]$ (which is a finite $\mathfrak{o}$-module). Let $\mathfrak{p}_1, \cdots, \mathfrak{p}_s$ be all the minimal prime divisors of $a\mathfrak{o}_1$. Then there exist a finite number of elements $c'_1, \cdots, c'_t$ in $\mathfrak{o}'$ such that $(\mathfrak{o}_1)_{\mathfrak{p}_i}[c'_1, \cdots, c'_t]$ is a normal ring for every i, by the corollary to Theorem 2. Set $\mathfrak{o}_2 = \mathfrak{o}_1[c'_1, \cdots, c'_t]$. Now we prove the following two lemmas:

LEMMA 1. *For any ring $\mathfrak{S}$ such that $\mathfrak{o}_2 \subseteq \mathfrak{S} \subseteq \mathfrak{o}'$ and for any prime ideal $\mathfrak{p}$ of rank 1 in $\mathfrak{S}$, the ring $\mathfrak{S}_{\mathfrak{p}}$ is a normal ring.*

Proof. If $a \notin \mathfrak{p}$, then $\mathfrak{S}_{\mathfrak{p}}$ contains $1/a$ and therefore $\mathfrak{S}_{\mathfrak{p}}$ is a ring of quotients of $\mathfrak{o}''$, which shows that $\mathfrak{S}_{\mathfrak{p}}$ is a normal ring in this case. When $\mathfrak{p}$ contains a, we set $\mathfrak{p}' = \mathfrak{p} \cap \mathfrak{o}_2$. Then $\mathfrak{p}'$ is of rank 1 by Theorem 1. Therefore $(\mathfrak{o}_2)_{\mathfrak{p}'}$ is a normal ring and it contains $\mathfrak{o}'$. Therefore $\mathfrak{S}_{\mathfrak{p}}$ is a ring of quotients of $\mathfrak{o}'$ and is a normal ring.

LEMMA 2. *With the same $\mathfrak{S}$ as in Lemma 1, let $\mathfrak{q}_1, \cdots, \mathfrak{q}_u$ be all the imbedded prime divisors of $a\mathfrak{S}$. If $d_1, \cdots, d_v$ are elements of $\mathfrak{o}'$ such that*

$\mathfrak{S}_{\mathfrak{q}_i}[d_1,\cdots,d_v]$ is a normal ring for every i and if $\mathfrak{q}'$ is an imbedded prime divisor of $a\mathfrak{S}'$, where $\mathfrak{S}'=\mathfrak{S}[d_1,\cdots,d_v]$, then $\mathfrak{q}'\cap\mathfrak{S}$ contains some $\mathfrak{q}_i$ properly, provided that $\mathfrak{S}$ is Noetherian.

Proof: Set $\mathfrak{q}=\mathfrak{q}'\cap\mathfrak{S}$. Since $\mathfrak{q}'$ is an imbedded prime divisor of $a\mathfrak{S}'$, $\mathfrak{S}'_{\mathfrak{q}'}$ is not a normal ring ([10, § 9]). Therefore $\mathfrak{S}_{\mathfrak{q}}$ is not normal because $\mathfrak{S}'\subseteq\mathfrak{o}'$. This shows that $\mathfrak{q}$ contains one of $\mathfrak{q}_i$, say $\mathfrak{q}_1$ by Lemma 1 ([10, § 9]). If $\mathfrak{q}=\mathfrak{q}_1$, we have a contradiction to the choice of $d_1,\cdots,d_v$ and our assertion is proved.

Now we proceed with the proof of Theorem 3. We choose elements $d_1,\cdots,d_v$ of $\mathfrak{o}'$ as in Lemma 2 with $\mathfrak{S}=\mathfrak{o}_2$; then repeat this process with $\mathfrak{S}=\mathfrak{o}_2[d_1,\cdots,d_v]$ and so on; the existence of such d_i's follows from the corollary to Theorem 2. Then by the finiteness of chains of prime ideals in $\mathfrak{o}_2$ (or by the finiteness of dimension over I of the field of quotients L of $\mathfrak{o}$), after a finite number of steps we reach a ring $\mathfrak{o}^*$ in which $a\mathfrak{o}^*$ has no imbedded prime divisor. Since $\mathfrak{o}_2[1/a]=\mathfrak{o}''$ is a normal ring, $\mathfrak{o}^*[1/a]$ is a normal ring. Therefore the fact that $a\mathfrak{o}^*$ has no imbedded prime divisor shows that $\mathfrak{o}^*$ is a normal ring (on account of the property proved in Lemma 1) ([10, § 9]). Therefore $\mathfrak{o}^*=\mathfrak{o}'$ and $\mathfrak{o}'$ is a finite $\mathfrak{o}$-module, because we added only a finite number of elements of $\mathfrak{o}'$ at each step.

6. Preliminaries from the theory of valuation rings.

LEMMA 1. *An integral domain $\mathfrak{v}$ is a valuation ring if and only if the set of all principal ideals of $\mathfrak{v}$ is linearly ordered by inclusion. Equivalently, if an element a of the field of quotients of $\mathfrak{v}$ is not in $\mathfrak{v}$, then a^{-1} is in $\mathfrak{v}$.* (Krull [4])

Since the proof is easy and is well known, we will omit it.

Remark. From this lemma, we see that if $\mathfrak{v}$ is a valuation ring and if $\mathfrak{p}$ is a prime ideal of $\mathfrak{v}$, then $\mathfrak{v}_{\mathfrak{p}}$ and $\mathfrak{v}/\mathfrak{p}$ are valuation rings. Further it is easy to see that $\mathfrak{p}\mathfrak{v}_{\mathfrak{p}}=\mathfrak{p}$ (set-theoretically).

LEMMA 2. *Let $\mathfrak{a}$ be an ideal of an integral domain $\mathfrak{o}$. Let L be a field containing $\mathfrak{o}$. If x is an element of L and if $\mathfrak{a}\mathfrak{o}[x]$ contains the identity, then $\mathfrak{a}\mathfrak{o}[x^{-1}]$ does not contain the identity.* (Chevalley; see Cohen-Seidenberg [3])

Proof. We may asume that $\mathfrak{a}$ is a prime ideal in $\mathfrak{o}$, for, if not, we may replace $\mathfrak{a}$ by any of its prime divisor. Similarly, considering $\mathfrak{o}_{\mathfrak{a}}$ instead of $\mathfrak{o}$, we may assume that $\mathfrak{a}$ is the unique maximal ideal of $\mathfrak{o}$. Since $1\in\mathfrak{a}\mathfrak{o}[x]$,

there exists a relation $1 = p_0 + p_1x + \cdots + p_nx^n$ with $p_i \in \mathfrak{a}$. Since $\mathfrak{a}$ is the unique maximal ideal of $\mathfrak{o}$, $1 - p_0$ is a unit in $\mathfrak{o}$ and x^{-1} is integral over $\mathfrak{o}$. Therefore $\mathfrak{a}\mathfrak{o}[x^{-1}] \neq \mathfrak{o}[x^{-1}]$ ([10, § 4]).

LEMMA 3. *Let $\mathfrak{o}$ be a subring of a field L and let $\mathfrak{p}$ be a prime ideal of $\mathfrak{o}$. Then there exists a valuation ring $\mathfrak{v}$ of L which dominates $\mathfrak{o}_{\mathfrak{p}}$.* (Krull [4])

Proof. Let F be the set of subrings $\mathfrak{s}$ of L which contain $\mathfrak{o}$ such that $\mathfrak{p}\mathfrak{s} \neq \mathfrak{s}$. Then the set F is an inductive set, because members $\mathfrak{s}$ of F are characterized by $1 \notin \mathfrak{p}\mathfrak{s}$ and $\mathfrak{o} \subseteq \mathfrak{s} \subseteq L$. Therefore there exists a maximal member $\mathfrak{v}$ in F by Zorn's lemma. If x is an element of L which is not in $\mathfrak{v}$, $\mathfrak{p}\mathfrak{v}[x]$ contains the identity by the maximality of $\mathfrak{v}$. Therefore $1 \notin \mathfrak{p}\mathfrak{v}[1/x]$ by Lemma 2 and we see that $1/x \in \mathfrak{v}$ again by the maximality of $\mathfrak{v}$. Therefore $\mathfrak{v}$ is a valuation ring by Lemma 1. The existence of $\mathfrak{v}$ which dominates $\mathfrak{o}_{\mathfrak{p}}$ is easily seen if we consider $\mathfrak{o}_{\mathfrak{p}}$ instead of $\mathfrak{o}$.

PROPOSITION 4. *Let $\mathfrak{v}$ be a valuation ring of a field L and let $\mathfrak{p}$ be the maximal ideal of $\mathfrak{v}$. Let $\mathfrak{o}$ be a valuation ring of the residue class field $\mathfrak{v}/\mathfrak{p}$. Then there exists a uniquely determined valuation ring $\mathfrak{v}'$ of L which contains $\mathfrak{p}$ as a prime ideal and satisfies $\mathfrak{v}'/\mathfrak{p} = \mathfrak{o}$. With this $\mathfrak{v}'$, we have $\mathfrak{v}'_{\mathfrak{p}} = \mathfrak{v}$.*

Proof. If there exists such $\mathfrak{v}'$, it must be the set of all elements of $\mathfrak{v}$ whose residue classes modulo $\mathfrak{p}$ are contained in $\mathfrak{o}$. Therefore we have only to show that this set $\mathfrak{v}'$ is a valuation ring, because if we see that $\mathfrak{v}'$ is a valuation ring then the last assertion is easy. Let a be an element of L which is not in $\mathfrak{v}'$. If a is not in $\mathfrak{v}$, then $1/a \in \mathfrak{p}$ and $1/a \in \mathfrak{v}'$. Assume that a is in $\mathfrak{v}$ and let a' be the residue class of a modulo $\mathfrak{p}$. Since $\mathfrak{o}$ is a valuation ring, $1/a'$ is in $\mathfrak{o}$ and $1/a \in \mathfrak{v}'$. Since obviously $\mathfrak{v}'$ is an integral domain, we see that $\mathfrak{v}'$ is a valuation ring.

We call $\mathfrak{v}'$ the *composite* of $\mathfrak{v}$ and $\mathfrak{o}$.

PROPOSITION 5. *Let $\mathfrak{o}$ be a subring of a field L and let $\mathfrak{p}_1 \subset \mathfrak{p}_2 \subset \cdots \subset \mathfrak{p}_r$ be a chain of prime ideals $\mathfrak{p}_i$ in $\mathfrak{o}$. Then there exists a valuation ring $\mathfrak{v}$ of L which has prime ideals $\mathfrak{n}_1, \cdots, \mathfrak{n}_r$ such that $\mathfrak{p}_i = \mathfrak{n}_i \cap \mathfrak{o}$ for each i.* (Krull [4])

Proof. We prove our assertion by induction on r. Let $\mathfrak{v}_1$ be a valuation ring of L whose maximal ideal $\mathfrak{n}_1$ lies over $\mathfrak{p}_1$; existence of $\mathfrak{v}_1$ follows from Lemma 3. Then by our induction asumption, there exists a valuation ring $\mathfrak{v}'$ of $\mathfrak{v}_1/\mathfrak{n}_1$ which has prime ideals $\mathfrak{n}'_2, \cdots, \mathfrak{n}'_r$ such that $\mathfrak{n}'_i \cap (\mathfrak{o}/\mathfrak{p}_1)$

$=\mathfrak{p}_i/\mathfrak{p}_1$ for each i. Then by Proposition 4, we see the existence of a valuation ring with the required property.

Remark. The proof of Proposition 5 shows that if we are given a valuation ring $\mathfrak{v}_1$ which dominates $\mathfrak{o}_{\mathfrak{p}_1}$, then the valuation ring $\mathfrak{v}$ may be chosen as the composite of $\mathfrak{v}_1$ and a valuation ring of the residue class field of $\mathfrak{v}_1$.

PROPOSITION 6. *Let $\mathfrak{o}$ be a subring of a field L. If an element x of L is not integral over $\mathfrak{o}$, then there exists a valuation ring $\mathfrak{v}$ of L which contains $\mathfrak{o}$ and does not contain x.* (Krull [4])

Proof. If x is integral over $\mathfrak{o}[1/x]$, then x is integral over $\mathfrak{o}$. Therefore x is not integral over $\mathfrak{o}[1/x]$ and we may assume that $1/x \in \mathfrak{o}$. Let $\mathfrak{p}$ be a prime ideal of $\mathfrak{o}$ containing $1/x$ and let $\mathfrak{v}$ be a valuation ring of L which dominates $\mathfrak{o}_{\mathfrak{p}}$ (by Lemma 3). Then $x^{-1}\mathfrak{v} \neq \mathfrak{v}$ and $\mathfrak{v}$ does not contain x.

COROLLARY. *A normal ring $\mathfrak{o}$ is the intersection of all valuation rings (of its field of quotients) which contain $\mathfrak{o}$ (and conversely).* (Krull [4])

Chapter 2. The Notion of Model.

Throughout this chapter, we fix a ground ring I unless the contrary is explicitly stated. Further, all function fields will be assumed to be contained in some fixed field, unless the contrary is explicitly stated.

1. Places and correspondences. A valuation ring (or a field) which is a ring of quotients of the ground ring I is called a *ground place* (of I). When L is a function field (over I), a valuation ring $\mathfrak{v}$ of L is called a *place* of L (over I) if $\mathfrak{v}$ dominates some ground place (of I). Since we fix the ground ring I, we shall often omit the term "over I" or "of I".

When P and P' are spots, we say that P and P' *correspond* to each other or that P corresponds to P' if there exists a place which dominates both P and P'. (The function fields of P and P' may be distinct from each other.)

THEOREM 1. *A spot P corresponds to a spot P' if and only if the ideal $\mathfrak{a}$ of $P[P']$ which is generated by the maximal ideals $\mathfrak{m}$ and $\mathfrak{m}'$ of P and P' does not contain 1. In that case, if $\mathfrak{p}$ is a prime ideal of $P[P']$ which contains $\mathfrak{a}$, then $P[P']_{\mathfrak{p}}$ is a spot which domiantes P, P'. Conversely, if a spot Q dominates both P and P', then Q dominates a spot which is a ring of quotients of $P[P']$ obtained as above.*

Proof. Assume that P corresponds to P'. Then there exists a place $\mathfrak{v}$ which dominates both P and P'. Let $\mathfrak{w}$ be the maximal ideal of $\mathfrak{v}$. Then $\mathfrak{w}$ contains $\mathfrak{m}$ and $\mathfrak{m}'$, and therefore also $\mathfrak{a}$, which shows that $\mathfrak{a}$ does not contain 1. Conversely, assume that $\mathfrak{a}$ does not contain 1. Let $\mathfrak{p}$ be a prime ideal of $P[P']$ which contains $\mathfrak{a}$. Let $\mathfrak{o}$ and $\mathfrak{o}'$ be affine rings such that P and P' are rings of quotients of $\mathfrak{o}$ and $\mathfrak{o}'$ respectively. Then $P[P']$ is a ring of quotients of $\mathfrak{o}[\mathfrak{o}']$ and $\mathfrak{o}[\mathfrak{o}']$ is an affine ring. Therefore $P[P']_{\mathfrak{p}}$ is a spot. Since $\mathfrak{p}$ lies over $\mathfrak{m}$ and $\mathfrak{m}'$ (because $\mathfrak{m}$ and $\mathfrak{m}'$ are maximal ideals), the maximal ideal $\mathfrak{p}P[P']_{\mathfrak{p}}$ of $P[P']_{\mathfrak{p}}$ lies over $\mathfrak{m}$ and $\mathfrak{m}'$, which shows that $P[P']_{\mathfrak{p}}$ dominates P and P'. Next we assume that a spot Q dominates both P and P'. Let $\mathfrak{v}'$ be any place which dominates Q. Then $\mathfrak{v}'$ dominates both P and P'. Therefore P and P' correspond to each other. Let $\mathfrak{p}'$ be the prime ideal of $P[P']$ over which the maximal ideal of Q lies. Then Q dominates $P[P']_{\mathfrak{p}'}$ and this last spot dominates P and P'. Thus the proof is completed.

LEMMA 1. *If two spots P and P' are rings of quotients of the same ring $\mathfrak{o}$, then P cannot correspond to P' unless $P = P'$.*

Proof. If $P \neq P'$, then there exists an element a of $\mathfrak{o}$ which is a unit in one of the rings P and P' and a non-unit in another, and P cannot correspond to P'.

2. Definition of models. A set A of spots of a function field L is called an *affine model* (of L) if there exists an affine ring $\mathfrak{o}$ of L such that A is the set of all spots which are rings of quotients of $\mathfrak{o}$ (with respect to prime ideals). Such an affine ring $\mathfrak{o}$ is called an affine ring of A and A is called the affine model defined by $\mathfrak{o}$. A given affine model can be defined by only one affine ring. Indeed:

LEMMA 1. *If A is an affine model and if $\mathfrak{o}$ is an affine ring which defines A, then $\mathfrak{o}$ is the intersection of all spots in A.*

This follows from the remark at the end of [10, § 9].

Now we define the notion of models. A non-empty set M of spots of a function field L is called a *model* of L if M is the union of a finite number of affine models of L and if two different spots in M never correspond to each other.

Observe that an affine model is a model by virtue of Lemma 2.1.1.

When M is a model of a function field L, we say that L is the function field of M.

Let M be a model of a function field L. When $\mathfrak{v}$ is a place of a function

field L' containing L, there may be a spot P in M which is dominated by $\mathfrak{v}$; if so, then P is uniquely determined; it is called the *center* of $\mathfrak{v}$ on M. A model M of a function field L is said to be *complete* if every place of L has a center on M.

Remark 1. An affine model is complete if and only if its affine ring is integral over the ground ring.

The proof follows from Proposition 1.6.

Remark 2. A model M of a function field L is complete if and only if one of the following conditions is satisfied:

(1) Every place of a function field containing L has a center on M.

(2) Every spot of L corresponds to a spot in M.

(3) Every spot of a function field which contains L corresponds to a spot in M.

In order to prove this, we first establish

LEMMA 2. *Let L be a function field and let L' be a function field which contains L. Then*

1) *If $\mathfrak{v}'$ is a place of L', then $\mathfrak{v}' \cap L$ is a place of L.*

2) *If $\mathfrak{v}$ is a place of L, then there exists a place $\mathfrak{v}'$ of L' such that $\mathfrak{v} = \mathfrak{v}' \cap L$.*

3) *If P is a spot of L, then there exists a place $\mathfrak{v}$ of L which dominates P.*

This follows immediately from results in § 6 of Chapter 1.

Proof of Remark 2. Equivalence with (1) is immediate from Lemma 2. We will prove the equivalence with (2) or (3). Let L' be any function field containing L. Assume first that M is complete. Let P be any spot of L or L'. Then there exists a place $\mathfrak{v}'$ of L' which dominates P. Let Q be the center of $\mathfrak{v}'$ on M. Then P corresponds to Q. Conversely, assume that M is not complete. Then there exists a place $\mathfrak{v}$ of L or L' which has no center on M. Let $A_1, \cdots, A_n$ be affine models of L such that $M = \bigcup_i A_i$ and let $\mathfrak{o}_i$ be the affine ring of A_i. Since $\mathfrak{v}$ has no center on M, $\mathfrak{v}$ does not contain any of the rings $\mathfrak{o}_i$; let x_i be an element of $\mathfrak{o}_i$ which is not in $\mathfrak{v}$ and set $y_i = 1/x_i$. Further let $y_{n+1}, \cdots, y_r$ be elements of $\mathfrak{v}$ such that $I[y_1, \cdots, y_r] = \mathfrak{S}$ is an affine ring of L or L'. Set $P = \mathfrak{S}_{(\mathfrak{m} \cap \mathfrak{S})}$, where $\mathfrak{m}$ is the maximal ideal of $\mathfrak{v}$. Let Q be any spot in M. Then Q is a ring of quotients of some $\mathfrak{o}_i$, say $\mathfrak{o}_1$.

Then x_1 is in Q and $y_1 = 1/x_1$ is a non-unit in P. Therefore y_1 is a unit in the ring $P[Q]$, which shows that P does not correspond to Q by Theorem 1. Thus we see that this spot P does not correspond to any spot in M and the proof is completed.

THEOREM 2. *Let $x_0 = 1, x_1, \cdots, x_n$ be elements of a function field L such that $I[x_1, \cdots, x_n]$ is an affine ring of L. Let A_i be the affine model defined by $\mathfrak{o}_i = I[x_0/x_i, \cdots, x_n/x_i]$ for each i such that $x_i \neq 0$. Then the union M of all A_i is a complete model of L.*

This model M is called the *projective model* of L defined by the *affine coordinates* $(x_1, \cdots, x_n)$. (A model is called a projective model if it is a projective model defined by a suitable affine coordinates.)

Proof. We first show that M is a model. Assume the contrary. Then there exist two different spots P and P' in M which correspond to each other. By Lemma 2.1.1, P and P' cannot be in the same affine model A_i. Since $(x_j/x_i)/(x_k/x_i) = x_j/x_k$, we may assume that P is in A_0 and P' is in A_1. Set $\mathfrak{o} = \mathfrak{o}_0[\mathfrak{o}_1]$. Then by Theorem 1, there exists a prime ideal $\mathfrak{p}$ of $\mathfrak{o}$ such that $\mathfrak{o}_\mathfrak{p}$ dominates both P and P'. Since $\mathfrak{o}$ contains both x_1 and $1/x_1$, x_1 is a unit in $\mathfrak{o}$. Since P contains x_1 and since $\mathfrak{o}_\mathfrak{p}$ dominates P, we see that x_1 is a unit in P and P contains $\mathfrak{o}$. It follows that $\mathfrak{o}_\mathfrak{p} = P$ (because P is a ring of quotients of $\mathfrak{o}_0$; see [10, § 2]). Similarly we see that $\mathfrak{o}_\mathfrak{p} = P'$, whence $P = P'$, which is a contradiction. Thus we have proved that M is a model. Now let $\mathfrak{v}$ be a place of L and let v be a valuation of L defined by $\mathfrak{v}$. We choose x_i such that $v(x_i) \leqq v(x_j)$ for any $j = 0, \cdots, n$. Then $\mathfrak{o}_i$ is contained in $\mathfrak{v}$. Let $\mathfrak{q}$ be the intersection of the maximal ideal of $\mathfrak{v}$ with $\mathfrak{o}_i$. Then $Q = (\mathfrak{o}_i)_\mathfrak{q}$ is dominated by $\mathfrak{v}$. Therefore Q is the center of $\mathfrak{v}$ on M, which shows that M is complete.

The notion of projective model can be defined by *homogeneous coordinates* as follows: Let $z_0, \cdots, z_n$ (at least one of z_i, say z_0, is not zero) be elements of a certain field containing I. Let A_i be the affine model defined by the affine ring $I[z_0/z_i, \cdots, z_n/z_i]$ for each i such that $z_i \neq 0$. Then the union of all A_i is the projective model defined by the affine coordinates $(z_1/z_0, \cdots, z_n/z_0)$. This model is called the projective model defined by the homogeneous coordinates $(z_0, \cdots, z_n)$. Observe that the projective model defined by the affine coordinates $(x_1, \cdots, x_n)$ is also defined by the homogeneous coordinates $(t, x_1t, \cdots, x_nt)$ with an arbitrary nonzero element t.

Let M be a projective model of a function field L. An affine ring $\mathfrak{o}$ is called a homogeneous coordinate ring of M if $\mathfrak{o}$ is generated by a system of homogeneous coordinates over the ground ring I and if it contains an

algebraically independent element over the function field L of M; when $\mathfrak{o}$ is expressed in the form $I[z_0, \cdots, z_n]$, then we understand that $(z_0, \cdots, z_n)$ is a system of homogeneous coordinates which defines M, unless the contrary is explicitly stated. Observe that one nonzero z_i is transcendental over L if and only if every nonzero z_j is transcendental over L. Let K be the field of quotients of $\mathfrak{o}$ and asume that z_0 is transcendental over L. Then $K = L(z_0)$. An element a of K is called homogeneous if there exists an element b of L such that $a = bz_0^r$ with a rational integer r; this r is called the degree of homogeneity (or merely degree) of a.

3. Specializations. A quasi-local ring P is said to be a *specialization* of another quasi-local ring P' if P' is a ring of quotients of P (necessarily with respect to a prime ideal). Observe that if a valuation ring $\mathfrak{v}$ is the composite of a valuation ring $\mathfrak{v}'$ and a valuation ring of the residue class field of $\mathfrak{v}'$, then $\mathfrak{v}$ is a specialization of $\mathfrak{v}'$.

PROPOSITION 1. *If a spot P' is a specialization of another spot P and if P' is in a model M, then P is also in M; if $P \neq P'$, then* $\dim P' < \dim P$.

Proof. Since there exists an affine model A such that $P' \varepsilon A \subseteq M$, we may assume that M is an affine model. Let $\mathfrak{o}$ be the affine ring of M and let $\mathfrak{p}'$ be the prime ideal of $\mathfrak{o}$ such that $P' = \mathfrak{o}_{\mathfrak{p}'}$. Let $\mathfrak{m}$ be the prime ideal of P' such that $P = P'_{\mathfrak{m}}$. Then $P = \mathfrak{o}_{(\mathfrak{m} \cap \mathfrak{o})}$, which shows that P is in M. If $P \neq P'$, then $\operatorname{rank} P < \operatorname{rank} P'$, whence $\dim P > \dim P'$ by Corollary 1 to Theorem 1.1.

PROPOSITION 2. *Assume that a spot P' of a function field L is a specialiaztion of another spot P (necessarily of L). If* $\dim P' < \dim P$, *then there exists a spot P^* of L such that* 1) *P' is a specialization of P^*,* 2) *P^* is a specialization of P and* 3) $\dim P^* = \dim P' + 1$.

The proof is easy making use of Corollary 1 to Theorem 1.1.

PROPOSITION 3. *Assume that the ground ring I is either a field or a Dedekind domain which has infinitely many prime ideal. If M is a model and if P is a spot in M, then there exists a spot P' in M which is a specialization of P and of dimension* 0.

Proof. We may assume that M is an affine model. Then the proof is immediate from Corollary 6 to Proposition 1.1.

PROPOSITION 4. *Let M be a model of a function field L and let L' be a function field which contains L. Assume that spots P and P' in M are*

dominated by places $\mathfrak{v}$ and $\mathfrak{v}'$ of L' respectively and that $\mathfrak{v}'$ is a specialization of $\mathfrak{v}$. Then P' is a specialization of P. Conversely, if a spot Q' of L is a specialization of another spot Q and if a place $\mathfrak{w}$ of L' dominates Q, then there exists a place $\mathfrak{w}'$ of L' which is a specialization of $\mathfrak{w}$ and dominates Q'.

Proof. Let A be an affine model of L which contains P' and is contained in M. Let $\mathfrak{o}$ be the affine ring of A and let $\mathfrak{m}$ and $\mathfrak{m}'$ be the maximal ideals of $\mathfrak{v}$ and $\mathfrak{v}'$ respectively. Then we have $\mathfrak{o} \subseteq P' \subseteq \mathfrak{v}' \subseteq \mathfrak{v}$. Therefore we can set $P^* = \mathfrak{o}_{(\mathfrak{m} \cap \mathfrak{o})}$ and $P^{**} = \mathfrak{o}_{(\mathfrak{m}' \cap \mathfrak{o})}$. Then P^* and P^{**} are the centers of $\mathfrak{v}$ and $\mathfrak{v}'$ on A, hence on M. Therefore $P^* = P$ and $P^{**} = P'$ and P' is a specialization of P because $\mathfrak{m} \subseteq \mathfrak{m}'$. The converse follows from the remark after Proposition 1.5.

4. Joins of models. Let M and M' be models (of function field L and L' respectively. We say that M *dominates* M' (in symbols $M' \leqq M$) if every spot in M dominates some spot in M'.

If P and P' are spots, the set of spots which are rings of quotients of $P[P']$ and dominate both P and P' is called the *join* of P and P'; it will be denoted by $J(P, P')$.

Remark. $J(P, P')$ is empty if and only if P does not correspond to P', by virtue of Theorem 1.

Now, for two sets M and M' of spots, the union of all $J(P, P')$, where P and P' run over all spots in M and M' respectively, is called the *join* of M and M' and will be denoted by $J(M, M')$.

THEOREM 3. *Let M and M' be models of function fields L and L' respectively. Then the join $J(M, M')$ of M and M' is a model of $L(L')$. Further $J(M, M')$ dominates both M and M' and if a model M'' (of a function field which contains $L(L')$) dominates M and M', it dominates $J(M, M')$.*

Proof. Our assertion, exception for the fact that $J(M, M')$ is a model, follows from Theorem 1. We prove that $J(M, M')$ is a model. Assume that spots P''_1 and P''_2 in $J(M, M')$ correspond to each other. We take spots $P_1, P_2 \in M$ and $P'_1, P'_2 \in M'$ such that $P''_i \in J(P_i, P'_i)$ for each i. Let $\mathfrak{v}$ be a place which dominates both P''_1 and P''_2. Then $\mathfrak{v}$ dominates P_1, P_2, P'_1 and P'_2. It follows that these spots correspond to each other. Therefore by the definition of models, we have $P_1 = P_2$ and $P'_1 = P'_2$. Thus we see that P''_1 and P''_2 are rings of quotients of the same ring; by Lemma 2.1.1, we have $P''_1 = P''_2$. Now we have only to show that $J(M, M')$ is the union of a finite number of affine models. In order to do this, we have only to prove

LEMMA 1. *If A and A' are affine models, then the join $J(A,A')$ is an affine model.*

Proof. Let $\mathfrak{o}$ and $\mathfrak{o}'$ be the affine rings of A and A' respectively. Then the ring $\mathfrak{o}''=\mathfrak{o}[\mathfrak{o}']$ is also an affine ring. Let A'' be the affine model defined by $\mathfrak{o}''$. By the definition of $J(A,A')$, every member of $J(A,A')$ is a ring of quotients of $\mathfrak{o}''$, whence $J(A,A')\subseteq A''$. Conversely, let P'' be a spot in A'' and let $\mathfrak{p}''$ be the prime ideal of $\mathfrak{o}''$ such that $P''=\mathfrak{o}''_{\mathfrak{p}''}$. Then P'' contains spots $P=\mathfrak{o}_{(\mathfrak{p}''\cap\mathfrak{o})}$ and $P'=\mathfrak{o}_{(\mathfrak{p}''\cap\mathfrak{o}')}$. Therefore P'' is a ring of quotients of $P[P']$ ([10, § 2]) and dominates P and P'. Therefore P'' is in $J(A,A')$ and $A''\subseteq J(A,A')$. Thus the assertion is proved.

PROPOSITION 5. *If M and M' are complete models of function fields L and L' respectively, then $J(M,M')$ is a complete model of $L(L')$. If M and M' are projective models, then so is $J(M,M')$.*

Proof. Let $\mathfrak{v}$ be a place of $L(L')$. Then $\mathfrak{v}$ has centers P and P' on M and M' respectively. Let $\mathfrak{m}$ be the maximal ideal of $\mathfrak{v}$ and set $\mathfrak{n}=\mathfrak{m}\cap P[P']$. Then $P''=P[P']_{\mathfrak{n}}$, which is in $J(M,M')$ by Theorem 1, is dominated by $\mathfrak{v}$, which shows that $\mathfrak{v}$ has a center P'' on $J(M,M')$. It follows that $J(M,M')$ is a complete model. Now we assume that M and M' are projective models defined by homogeneous coordinates $(x_0,\cdots,x_m)$ and $(y_0,\cdots,y_n)$ respectively. We may assume here that no x_i and no y_j is zero. Let M'' be the projective model defined by the homogeneous coordinates $(x_iy_j;i=0,\cdots,m;j=0,\cdots,n)$. We denote by $\mathfrak{o}_i$ and $\mathfrak{o}'_j$ the affine rings $I[x_0/x_i,\cdots,x_m/x_i]$ and $I[y_0/y_j,\cdots,y_n/y_j]$ respectively. Let A_i and A'_j be the affine models defined by $\mathfrak{o}_i$ and $\mathfrak{o}'_j$ respectively. On the other hand, let $\mathfrak{o}''_{ij}$ be the affine ring $\mathfrak{o}_i[\mathfrak{o}'_j]$ and let A''_{ij} be the affine model defined by $\mathfrak{o}''_{ij}$. Then $\mathfrak{o}''_{ij}$ is generated by the elements x_ry_s/x_iy_j (i and j are fixed; $0\leqq r\leqq m$, $0\leqq s\leqq n$) over I. Therefore M'' is the union of all A''_{ij}. Since $\mathfrak{o}''_{ij}=\mathfrak{o}_i[\mathfrak{o}'_j]$, $A''_{ij}=J(A_i,A'_j)$. Therefore M'' is the join $J(M,M')$ of M and M'. Thus we see that $J(M,M')$ is a projective model.

If a model M dominates another model M', then $J(M,M')=M$. From this fact we deduce

LEMMA 2. *Assume that a model M of a function field L dominates a model M' of a function field L'. Then there exist affine models A_{ij} of L and affine models A'_j of L' $(j=1,\cdots,m;i=1,\cdots,n(j))$ such that 1) $A_{ij}\geqq A'_j$ and 2) $M=\bigcup_{ij}A_{ij}$, $M'=\bigcup_j A'_j$.*

LEMMA 3. *Let M and M' be models of the same function field L.*

Assume that there exists a model M'' of L which contains both M and M', then $M \cap M' = J(M, M')$.

The proof is easy.

5. Derived normal model of a model. We say that a model M is *normal* if every spot in M is normal.

NOTATION. *Let P be a spot of a function field L and let L' be a finite algebraic extension of L. Further let $\mathfrak{o}$ be the integral closure of P in L'. Then we denote by $N(P;L')$ the set of spots which are rings of quotients of $\mathfrak{o}$ with respect to maximal ideals of $\mathfrak{o}$. When M is a set of spots of L, then the union of all $N(P;L')$, where P runs over all spots in M, will be denoted by $N(M;L')$. On the other hand, $N(P;L)$ and $N(M;L)$ are denoted by $N(P)$ and $N(M)$ respectively.*

Remark 1. $N(P;L')$ is a finite set, because $\mathfrak{o}$ is a finite P-module.[3]

Remark 2. In §§ 1-4, our assumption that ground rings satisfy the finiteness condition for integral extensions did not play any rôle. In the present section this condition will be used in an esential manner. Without it, it would not be true in general that the derived normal ring of a spot P is a finite P-module. (If the function field is separably generated over the ground ring I, then the derived normal ring of P is a finite P-module; cf. the appendix to the second paper of this series; in this case, assuming further that L' is separably generated, the results of the present section hold without assuming the finiteness condition.)

THEOREM 4. *Let M be a model of a function field L and let L' be a finite algebraic extension of L. Then $N(M;L')$ is a normal model. If a normal model M' of a function field containing L' dominates M, then M' dominates $N(M;L')$.*

This model $N(M;L')$ is called the *derived normal model* of M in L'; $N(M)$ is called the derived normal model of M.

Proof. We first prove the first assertion in the case where M is an affine model. Let $\mathfrak{o}$ be the affine ring of M and let $\mathfrak{o}'$ be the integral closure of $\mathfrak{o}$ in L'. Then by Theorem 1.3, $\mathfrak{o}'$ is also an affine ring. On the other hand, $N(M;L')$ is the set of spots which are rings of quotients of $\mathfrak{o}'$, which shows

[3] It is well known that if $\mathfrak{o}$ is a (Noetherian) semi-local integral domain, then the derived normal ring of $\mathfrak{o}$ has only a finite number of maximal ideals (but it may not be Noetherian).

that $N(M;L')$ is an affine model and is normal. Now we consider the general case. By the above statement, we see that $N(M;L')$ is the union of a finite number of affine models. Assume that a spot $P^* \varepsilon N(M;L')$ corresponds to a spot $P^{**} \varepsilon N(M;L')$. Let P and P' be spots in M such that $P^* \varepsilon N(P;L')$ and $P^{**} \varepsilon N(P';L')$. Let $\mathfrak{v}$ be a place which dominates both P^* and P^{**}. Since P^* and P^{**} dominate P and P' respectively, $\mathfrak{v}$ dominates both P and P', which shows that $P=P'$. Therefore P^* and P^{**} are rings of quotients of the same ring (which is the integral closure of P in L'). Therefore by Lemma 2.1.1, we have $P^*=P^{**}$. Thus we see that $N(M;L')$ is a model and is obviously normal. Now we prove the last assertion. Let P'' be a spot in M'. Then there exists a spot P in M which is dominated by P''. Since P'' is a normal ring and since the field of quotients of P'' contains L', P'' contains the integral closure $\mathfrak{o}'$ of P in L'. Let $\mathfrak{m}''$ be the maximal ideal of P''. Since P'' dominates P, $\mathfrak{m}'' \cap P$ is the maximal ideal of P. Therefore $\mathfrak{m}'' \cap \mathfrak{o}'$ lies over the maximal ideal of P, which shows that $\mathfrak{m}'' \cap \mathfrak{o}'$ is a maximal ideal of $\mathfrak{o}'$ ([10, § 4]). Therefore the ring $\mathfrak{o}'_{(\mathfrak{m}'' \cap \mathfrak{o}')}$ is in $N(M;L')$ and is dominated by P''. This completes the proof.

THEOREM 5. *Let M be a model of a function field L and let L' be a finite algebraic extension of L. If M is a complete model, then $N(M;L')$ is also complete. If M is a projective model, then $N(M;L')$ is also a projective model.*

Proof. Since the first half is easy, we prove only the last assertion. We assume that M is the projective model defined by the homogeneous coordinates $(y_0,\cdots,y_n)$, where we may assume that y_0 is transcendental over L and that all y_i are not zero. Set $\mathfrak{h}=I[y_0,\cdots,y_n]$. Since $\mathfrak{h}$ is an affine ring, the integral closure $\mathfrak{h}'$ of $\mathfrak{h}$ in $L'(y_0)$ is a finite $\mathfrak{h}$-module. Let $\mathfrak{i}$ be the set of elements of $\mathfrak{h}'$ which are homogeneous and of positive degree (where the notion of homogeneity and of degree are defined in the same way as in the case of $L(y_0)$). Since $\mathfrak{h}'$ is a finite $\mathfrak{h}$-module, $\mathfrak{h}''=\mathfrak{h}[\mathfrak{i}]$ is a finite $\mathfrak{h}$-module. Let r be a natural number and let $w_0,\cdots,w_t$ be a set of generators of the I-module of all elements of $\mathfrak{h}''$ of degree r; here we choose w_i to be y_i^r for $i \leqq n$. Further, let $M(r)$ be the projective model defined by the homogeneous coordinates $(w_0,\cdots,w_t)$ and let $L(r)$ be the function field of $M(r)$. It is obvious that $L(r)$ is contained in L'. We set $z_{ij}=w_i/w_j$ and $\mathfrak{o}_i=I[z_{0i},\cdots,z_{ti}]$. Then

(1) For each $i \leqq n$, $\mathfrak{o}_i$ is integral over $I[y_0/y_i,\cdots,y_n/y_i]$.

Proof. Since w_j is integral over $\mathfrak{h}$, there exists a relation $w_j^u+c_1w_j^{u-1}$

$+\cdots+c_u=0$ with $c_k \varepsilon \mathfrak{h}$; here we may assume that each c_k is a homogeneous polynomial in $y_0,\cdots,y_n$ of degree rk. Then c_k/y_i^{rk} is in $I[y_0/y_i,\cdots,y_n/y_i]$ and w_j/y_i^r is integral over $I[y_0/y_i,\cdots,y_n/y_i]$.

(2) For sufficiently large r and for $i \leqq n$, $\mathfrak{o}_i$ is the integral closure of $I[y_0/y_i,\cdots,y_n/y_i]$ in L'.

Proof. Let c be any element of the integral closure $\mathfrak{o}'_i$ of $I[y_0/y_i,$ $\cdots,y_n/y_i]$ in L'. Then there exists one integer d such that cy_i^d is integral over $\mathfrak{h}$. If r is chosen to be not less than d, then, since cy_i^r is also in $\mathfrak{h}''$ in this case, we see that c is in $\mathfrak{o}_i$. Since $\mathfrak{o}'_i$ is a finite $I[y_0/y_i,\cdots,y_n/y_i]$-module, we see that for sufficiently large r, $\mathfrak{o}_i$ contains $\mathfrak{o}'_i$. Now, by virtue of the observation in (1) above, we see that $\mathfrak{o}_i=\mathfrak{o}'_i$ for sufficiently large r.

The fact proved in (2) shows that $N(M;L')$ is a subset of $M(r)$ for sufficiently large r. Since $N(M;L')$ is a complete model, it follows that $N(M;L')=M(r)$ and $N(M;L')$ is a projective model of L'. Thus the assertion is proved.

6. Irreducible sets. When M is a model and when P is a spot in M, the set of spots in M which are specialization of P is called the *locus* of P in M and will be denoted by $M(P)$.

A subset E of a model M is called an *irreducible set* if there exists a spot P in M which satisfies the following two conditions: (1) E is contained in the locus of P in M and (2) for any finite number of spots $P_1,\cdots,P_n$ in M which are specializations of P and which are different from P, E is not contained in the union of the loci $M(P_i)$ of the spots P_i. We call P the *generating spot* of E; it is unique as we now show.

Assume that P' is also a generating spot of E. Let Q be a spot in E and let A be an affine model containing Q and contained in M. By Proposition 1 P and P' are in A. Let $\mathfrak{o}$ be the affine ring of A and let $\mathfrak{p}$ and $\mathfrak{p}'$ be the prime ideals of $\mathfrak{o}$ such that $P=\mathfrak{o}_{\mathfrak{p}}$, $P'=\mathfrak{o}_{\mathfrak{p}'}$. Let $\mathfrak{p}_1,\cdots,\mathfrak{p}_r$ be the set of all minimal prime divisors of $\mathfrak{p}+\mathfrak{p}'$ and set $P_i=\mathfrak{o}_{\mathfrak{p}_i}$. Since Q is a specialization of both P and P', the prime ideal $\mathfrak{q}$ such that $Q=\mathfrak{o}_{\mathfrak{q}}$ contains $\mathfrak{p}+\mathfrak{p}'$. Therefore Q is a specialization of some P_i. This shows that $E\cap A$ is contained in the union of the loci of the spots $P_1,\cdots,P_r$. Since the same holds for any affine model contained in M and since M is the union of a finite number of affine models, we have a contradiction if $P\neq P'$.

Remark. Since the notion of specialization does not depend on the choice of the model which carries the given spots, the notion of irreducible set does not depend on the choice of the model which carries the given sets.

LEMMA 1. *Let A be the affine model of a function field L defined by an affine ring $\mathfrak{o}$. Let E be a subset of A. Then E is irreducible if and only if the intersection $\mathfrak{p}$ of all prime ideals $\mathfrak{q}$ such that $\mathfrak{o}_{\mathfrak{q}}$ is in E is a prime ideal. In this case, $\mathfrak{o}_{\mathfrak{p}}$ is the generating spot of E.*

Proof. Assume first that $\mathfrak{p}$ is a prime ideal. If $\mathfrak{o}_{\mathfrak{q}}$ is in E ($\mathfrak{q}$ being a prime ideal of $\mathfrak{o}$), $\mathfrak{q}$ contains $\mathfrak{p}$ and $\mathfrak{o}_{\mathfrak{q}}$ is a specialization of the spot $P = \mathfrak{o}_{\mathfrak{p}}$. If there exist spots $\mathfrak{o}_{\mathfrak{p}_1}, \cdots, \mathfrak{o}_{\mathfrak{p}_n}$ ($\mathfrak{p}_i$ being prime ideals of $\mathfrak{o}$) which are specializations of P and are different from P such that E is contained in the union of $A(\mathfrak{o}_{\mathfrak{p}_i})$, then by the definition of $\mathfrak{p}$, the intersection of all $\mathfrak{p}_i$ must be contained in $\mathfrak{p}$, which is a contradiction because $\mathfrak{p} \subset \mathfrak{p}_i$ ($\mathfrak{p} \neq \mathfrak{p}_i$). This shows that P is the generating spot of E. Conversely, assume that $\mathfrak{p}$ is not prime. Since $\mathfrak{p}$ is an intersection of prime ideals, $\mathfrak{p}$ is a semi-prime ideal. Therefore $\mathfrak{p}$ is the intersection of the minimal prime divisors $\mathfrak{p}_1, \cdots, \mathfrak{p}_r$ ($r > 1$) of $\mathfrak{p}$ ([10, § 1]). Set $P_i = \mathfrak{o}_{\mathfrak{p}_i}$. If P is a spot whose locus contain E, then the prime ideal $\mathfrak{q}$ such that $P = \mathfrak{o}_{\mathfrak{q}}$ must be contained in $\mathfrak{p}$. Since E is contained in the union of the loci $P_1, \cdots, P_r$, P is not a generating spot of E, which shows that E is not irreducible. Thus the proof is completed.

On the other hand, with the same notations as above, if we set $E_i = E \cap A(P_i)$ for each i, then E_i is an irreducible set with generating spot P_i, which shows that E is the union of the irreducible sets E_i. Since a model is the union of a finite number of affine models, we have

LEMMA 2. *Every subset E of a model M is the union of a finite number of irreducible sets.*

COROLLARY. *If a subset E of a model M contains infinitely many spots, then there exists an irreducible subset of E which contains infinitely many spots.*

Again by the fact that a model is the union of a finite number of affine models, we see easily

LEMMA 3. *Let E be a subset of a model M of a function field L. Then a spot $P \in M$ is the generating spot of E (and consequently E is irreducible) if and only if* 1) *E is contained in the locus of P in M and* 2) *there exists an affine model A of L which is contained in M such that P is the generating spot of $E \cap A$.*

Remark. Let E be a subset of a model M. Then there may be different expressions of E as the union of a finite number of irreducible sets. But the following fact holds good.

There are irreducible sets $E_1, \cdots, E_n$ which satisfy the following three conditions; these E_i are uniquely determined:

1) E is the union of all E_i. 2) Every irreducible subset of E is contained in at least one of E_i. 3) There is no inclusion relation among the E_i.

Proof. Let $F_1, \cdots, F_r$ be irreducible subsets of E such that E is the union of them. Let P_i be the generating spot of F_i and set $E_i = E \cap M(P_i)$. We may assume without loss of generality that there is no inclusion relation among $E_1, \cdots, E_n$ and that each of $E_{n+1}, \cdots, E_r$ is contained in some E_i $(i \leqq n)$. These $E_1, \cdots, E_n$ satisfy the conditions 1) and 3). Let F be an irreducible subset of W with generating spot Q. Let A be an affine model contained in M such that $F \cap A$ is an irreducible set with generating spot Q (by Lemma 3). Then by the proof of Lemma 1, we see that Q is a specialization of some P_i, which shows that F is contained in some E_i. Therefore these E_i satisfy the condition 2). If $E'_1, \cdots, E'_m$ satisfy the above three conditions, then each E_i must be contained in some E'_j and E'_j must be contained in some E_k. Therefore the systems $E_1, \cdots, E_n$ and $E'_1, \cdots, E'_n$ must coincide with each other.

7. Zariski topology. Let M be a model. Let $\mathfrak{F}$ be the family of subsets F of M which satisfy the following conditions: (1) If a spot P is in F, then F contains the locus of P in M and (2) if F contains an irreducible set E, then F contains the generating spot of E.

This $\mathfrak{F}$ can be used as the family of closed sets in a topology, called the *Zariski topology* in M.

THEOREM 6. *A subset F of a model M is a closed set of M if and only if there exist a finite number of spots $P_1, \cdots, P_n$ in M such that F is the union of the loci $M(P_i)$.*

Proof. Each $M(P_i)$ is a closed set and therefore $\bigcup_i M(P_i)$ is also a closed set. Assume that F is a closed set of M. Let F' be the set of spots P in F which are not specialization of any other spots in F. We have only to show that F' is a finite set, which follows immediately from Lemma 2.6.2.

PROPOSITION 6. *A subset F of a model M is a closed set of M if and only if the following two conditions are satisfied*: (1) *If P is a spot in F, then $M(P) \subseteq F$ and* (2) *if a spot $P' \in M$ is not in F, then there exist a finite number of spots $P_1, \cdots, P_n \in M(P')$ which are different from P' such that $F \cap M(P') \subseteq \bigcup_i M(P_i)$.*

Proof. Assume first that F is closed. By the definition of closed sets, condition (1) must hold. Since F is a closed set, $F \cap M(P')$ is also a closed set. Therefore by Theorem 6, condition (2) holds also. We shall now prove the converse part. Let F' be defined as in the proof of Theorem 6. Then we have only to show that F' is a finite set. Assume the contrary. Then there exists an irreducible set F'' which contains infinitely many spots and is contained in F' by the corollary to Lemma 2. 6. 2. Let P'' be the generating spot of F''. Then by our assumption, P'' is not in F and therefore there exist a finite number of spots $P_1, \cdots, P_n \varepsilon M(P'')$ $(P_i \neq P'')$ such that F'' is contained in $\bigcup_i M(P_i)$ by condition (2), which is a contradiction to our assumption that F'' is irreducible. Thus the proof is completed.

THEOREM 7. *Let A be the affine model defined by an affine ring $\mathfrak{o}$. Then a subset F of A is a closed set of A if and only if there exists an ideal $\mathfrak{a}$ of $\mathfrak{o}$ such that F is the set of spots which are rings of quotients of $\mathfrak{o}$ with respect to prime ideals which contain $\mathfrak{a}$.*

In this case, we say that F is the *closed set of A defined by the ideal* $\mathfrak{a}$ and that $\mathfrak{a}$ is an ideal of $\mathfrak{o}$ which defines F.

Proof. If F is a closed set, then by Theorem 6, there exist a finite number of spots $P_1, \cdots, P_n$ such that $F = \bigcup_i A(P_i)$. Let $\mathfrak{p}_i$ be the prime ideal of $\mathfrak{o}$ such that $P_i = \mathfrak{o}_{\mathfrak{p}_i}$. Then $\mathfrak{a} = \bigcap_i \mathfrak{p}_i$ has the required property. Conversely, asume that $\mathfrak{a}$ is an ideal of $\mathfrak{o}$. Let $\mathfrak{p}_1, \cdots, \mathfrak{p}_n$ be the minimal prime divisors of $\mathfrak{a}$. Then a prime ideal $\mathfrak{p}$ of $\mathfrak{o}$ contains $\mathfrak{a}$ if and only if $\mathfrak{p}$ contains one of $\mathfrak{p}_i$'s. It follows that the set of rings of quotients of $\mathfrak{o}$ with respect to prime ideals containing $\mathfrak{a}$ is a closed set. Thus the proof is completed.

THEOREM 8. *Let M be a model of a function field L and let $A_1, \cdots, A_n$ be models of L such that M is the union of them. Then a subset F of M is closed if $F \cap A_i$ is a closed set of A_i for each i. Conversely, if F is a closed set of M, then $F \cap A$ is a closed set of A for any model A of L contained in M.*

Proof. Assume that $F \cap A_i$ is closed in A_i for each i. Then there exist a finite number of spots P_{ij} $(i = 1, \cdots, n; j = 1, \cdots, m(i))$ such that $F \cap A_i = \bigcup_j A_i(P_{ij})$. Then $F = \bigcup_{ij} A_i(P_{ij}) \subseteq \bigcup_{ij} M(P_{ij})$. We denote this last set by F'. Let P' be a spot in F'. Let i be such that P' is in A_i. Then $P' \varepsilon \bigcup_j A_i(P_{ij})$, which shows that $F = F'$ and we see that F is a closed set of M. Conversely, assume that F is a closed set of M. Take spots $P_1, \cdots, P_n \varepsilon F$ such that $F = \bigcup_i M(P_i)$. If $M(P_i)$ meets A, then P_i is in

A by Proposition 1 and in this case $M(P_i) \cap A$ coincides with $A(P_i)$. Therefore we see that $F \cap A$ is a closed set of A.

COROLLARY. *Let M' be a model of a function field L. If M' is contained in a model M of L, then M' is a subspace of M.*

Proof. If F is a closed set of M, then $F \cap M'$ is a closed set of M' by Theorem 8. Assume that F' is a closed set of M'. Take spots $P_1, \cdots, P_n \varepsilon F'$ such that $F' = \bigcup_i M'(P_i)$ and set $F = \bigcup_i M(P_i)$. Then F is a closed set of M and $F \cap M' = F'$. Thus the corollary is proved.

THEOREM 9. *Let M be a model of a function field L. A non-empty subset M' of M is again a model of L if and only if M' is an open set of M.*

Proof. Assume that M' is a model. We may assume that M is an affine model; because for every affine model A contained in M, $M' \cap A = J(M', A)$ (by Lemma 2.4.3) is a model and therefore, if our theorem is proved for $M = A$, then Theorem 8 asserts that M' is open. Further, we may assume that M' is an affine model, because M' is the union of a finite number of affine models. Let $\mathfrak{o}$ and $\mathfrak{o}'$ be affine rings of M and M' respectively and let F be the complements of M' in M. We have only to show that F is a closed set of M. Since M contains M', $\mathfrak{o}'$ contains $\mathfrak{o}$. Let $a_1, \cdots, a_n$ be elements of $\mathfrak{o}'$ such that $\mathfrak{o}' = \mathfrak{o}[a_1, \cdots, a_n]$ and set $\mathfrak{a}_i = \{a; a \varepsilon \mathfrak{o}, aa_i \varepsilon \mathfrak{o}\}$. Then each $\mathfrak{a}_i$ is an ideal of $\mathfrak{o}$. Set $\mathfrak{a} = \bigcap_i \mathfrak{a}_i$ and let F' be the closed set of M defined by $\mathfrak{a}$. Then it is sufficient to show that $F = F'$. Let $\mathfrak{p}$ be a prime ideal of $\mathfrak{o}$. Then $\mathfrak{o}_\mathfrak{p} \varepsilon F'$ if and only if $\mathfrak{a} \subseteq \mathfrak{p}$. Equivalently, $\mathfrak{p}$ contains some $\mathfrak{a}_i$, which means $\mathfrak{o}_\mathfrak{p}$ does not contain a_i, i.e., $\mathfrak{o}_\mathfrak{p}$ is not in M', $\mathfrak{o}_\mathfrak{p}$ is in F. Thus we see that $F = F'$. Now we shall prove the converse part of our theorem. Assume that M' is an open set of M. Let A_i $(i = 1, \cdots, n)$ be affine models such that M is the union of them. Then each A_i is an open set by the above proof. Therefore $M'_i = M' \cap A_i$ is an open set of M and therefore of A_i. If we can show that M'_i is a model, it is the union of a finite number of affine models, and the same will be true of M' and M' will be a model. Thus it suffices to prove the theorem in the case $M = A_i$ is an affine model. Let $\mathfrak{o}$ be the affine ring of M and let F be the complements of M' in M. Then there exists an ideal $\mathfrak{a}$ of $\mathfrak{o}$ which defines F by Theorem 7. Let $a_1, \cdots, a_r$ be nonzero elements of $\mathfrak{o}$ which generate $\mathfrak{a}$ (of $\mathfrak{a} = 0$, then $F = M$ and M' is empty, which is not the case). Let A'_j be the affine model defined by $\mathfrak{o}'_j = \mathfrak{o}[a_j^{-1}]$ for each j and set $M'' = \bigcup_j A'_j$. Each set A'_j being contained in M, M'' is contained in M and M'' is a model. Therefore it is sufficient to show that M' coincides with M''. Let $\mathfrak{p}$ be a prime ideal of $\mathfrak{o}$. Then $\mathfrak{o}_\mathfrak{p} \varepsilon M'$ if and

only if $\mathfrak{p}$ does not contain $\mathfrak{a}$. Equivalently, there exists one i such that $a_i \notin \mathfrak{p}$, i.e., $a_i^{-1} \varepsilon \mathfrak{o}_{\mathfrak{p}}$ and $\mathfrak{o}_{\mathfrak{p}} \varepsilon A'_i$, which is equivalent to say that $\mathfrak{o}_{\mathfrak{p}} \varepsilon M''$. Thus $M' = M''$ and the assertion is proved.

LEMMA 1. *Let A and A' be affine models of the same function field L. Assume that A dominates A'. Then $A \cap A'$ is a model.*

Proof. Let $\mathfrak{o}$ and $\mathfrak{o}'$ be the affine rings of A and A' respectively. Since A dominates A', $\mathfrak{o}$ contains $\mathfrak{o}'$. Take elements $a_1, \cdots, a_n$ of $\mathfrak{o}$ such that $\mathfrak{o} = \mathfrak{o}'[a_1, \cdots, a_n]$. We set $\mathfrak{a} = \{a ; a \varepsilon \mathfrak{o}', aa_i \varepsilon \mathfrak{o}' \text{ for every } i\}$. Then $\mathfrak{a}$ is an ideal of $\mathfrak{o}'$. Let F be the closed set of A' defined by $\mathfrak{a}$. It is sufficient to show that F coincides with the complements F' of $A \cap A'$ in A' by virtue of Theorem 9 (observe that $A \cap A'$ is not empty because L is in every model of L). Let $\mathfrak{p}'$ be a prime ideal of $\mathfrak{o}'$. Then $\mathfrak{o}'_{\mathfrak{p}'} \varepsilon F$ if and only if $\mathfrak{a} \subseteq \mathfrak{p}'$. Equivalently, there exists one i such that $a_i \notin \mathfrak{o}'_{\mathfrak{p}'}$ and $\mathfrak{o}'_{\mathfrak{p}'} \varepsilon F'$. Thus $F = F'$ and the lemma is proved.

THEOREM 10. *Let M and M' be models of the same function field L. Then $M \cap M'$ is also a model of L.*

Proof. Since $M \cap M' = J(M, M') \cap M'$, we may assume without loss of generality that M dominates M'. Let A_{ij} and A'_i $(i = 1, \cdots, m; j = 1, \cdots, n(i))$ be affine models of L such that $M = \bigcup_{ij} A_{ij}$, $M' = \bigcup_i A'_i$ and $A_{ij} \geqq A'_i$ for every i and j (Lemma 2.4.2). Then by the preceding lemma, we see that $A_{ij} \cap A'_i$ is a model, which shows that $M'' = \bigcup_{ij} (A_{ij} \cap A'_i)$ is a model (because it is a subset of a model and is the union of a finite number of affine models). Therefore it is sufficient to show that M'' coincides with $M \cap M'$. Assume that $P \varepsilon M \cap M'$. Then P is in some A_{ij}. Since A_{ij} dominates A'_i, A'_i contains a spot P' which is dominated by P. Since P is in M', we have $P' = P$ and $P \varepsilon A_{ij} \cap A'_i \subseteq M''$. Thus $M \cap M' \subseteq M''$. Since the converse inclusion is obvious, we have $M \cap M' = M''$ and $M \cap M'$ is a model.

COROLLARY 1. *Let M be a model. Then the set M' of normal spots in M is a model.*

The proof is easy if we observe that $M' = M \cap N(M)$.

COROLLARY 2. *Let M and M' be models of function fields L and L' respectively. Assume that L contains L'. Then the set M'' of spots in M which dominate spots in M' is a model.*

The proof is easy if we observe that $M'' = M \cap J(M, M')$.

Now we shall add some remarks on closed sets.

1) If F is a closed set of a model M, F is the union of a finite number of closed sets $F_1, \cdots, F_n$ of M which are irreducible. Here, if there is no inclusion relation $F_1, \cdots, F_n$, we say that each F_i is an *irreducible component* of F. The set of irreducible components of a closed set is uniquely determined. (The proof is similar to that of the remark at the end of § 6.)

2) If E is an irreducible set in a model M and if P is the generating spot of E, then the closure of E in M is the locus of P in M, which shows that it is an irreducible set. (The proof is easy.)

3) With the same notations as in the remark at the end of § 6, the closure of E is the union of the closures of $E_1, \cdots, E_n$. Further, each of the closure of $E_1, \cdots, E_n$ is an irreducible component of the closure of E. (The proof is easy by virtue of the result in 2) above.)

4) By virtue of Proposition 6, we see that the minimum condition for closed sets holds in a model. Consequently, a model is a compact space.

8. Induced models, reduced models and local models. Let P be a quasi-local ring with maximal ideal $\mathfrak{m}$. Then the natural homomorphism from P onto $P/\mathfrak{m}$ is called the *homomorphism defined by* P; it will be denoted by ϕ_P. If M is a set of rings contained in P, the set of homomorphic images of rings in M under ϕ_P (which will be called merely the image of M by ϕ_P) will be denoted by $\phi_P(M)$.

Let M be a model of a function field L over a ground ring I.

PROPOSITION 7. *Let F be an irreducible closed set of M with generating spot P. Then $\phi_P(F)$ is a model of $\phi_P(P)$ over $\phi_P(I)$.*

This model $\phi_P(F)$ is called the *induced model* defined by F, or the induced model of M defined by P; this may be denoted also by $\phi_P(M)$ (because this is uniquely determined by M and P).

Proof. If M is the affine model defined by an affine ring $\mathfrak{o}$, then $\phi_P(F)$ is the affine model defined by $\phi_P(\mathfrak{o})$ over $\phi_P(I)$. From this, in the general case, it follows that $\phi_P(F)$ is the union of a finite number of affine models over $\phi_P(I)$. Assume that two spots $\phi_P(Q)$ and $\phi_P(R)$ in $\phi_P(F)$ $(Q, R \varepsilon F)$ correspond to each other. Then there exists a place $\mathfrak{v}'$ of $\phi_P(P)$ over $\phi_P(I)$ which dominates both $\phi_P(Q)$ and $\phi_P(R)$. Let $\mathfrak{v}$ be a place of L which dominates P and let $\mathfrak{o}$ be a valuation ring of the residue class field of $\mathfrak{v}$ which dominates $\mathfrak{v}'$. Then the composite $\mathfrak{v}''$ of $\mathfrak{v}$ and $\mathfrak{o}$ dominates both Q and R;

Q corresponds to R, whence $Q=R$ and $\phi_P(Q)=\phi_P(R)$. This completes our proof.

Remark. The above proof shows also that spots in F are mapped in a one-to-one way onto spots in $\phi_P(F)$.

Let M be a model over a ground ring I and let $\mathfrak{p}$ be a prime ideal of I. A spot P in M is called a general spot over $\mathfrak{p}$ (or over the ground place $I_{\mathfrak{p}}$) if 1) P dominates $I_{\mathfrak{p}}$ and 2) P is of rank 0 or 1 according as $\mathfrak{p}=0$ or not. (If $\mathfrak{p}=0$, then P is just the function field of M.) If, for a given prime ideal $\mathfrak{p}$ of I, there exists one and only one general spot P over $\mathfrak{p}$, then the induced model $\phi_P(M)$ is called the *induced model of M modulo $\mathfrak{p}$*.

Let I' be a subring of L such that 1) there exists an affine ring $\mathfrak{S}$ over I such that I' is a ring of quotients of $\mathfrak{S}$ and 2) I' is a Dedekind domain. Since $\mathfrak{S}$ satisfies the finiteness condition for integral extensions by Theorem 1.3, I' also does and I' is a ground ring.

PROPOSITION 8. *Let $M_{I'}$ be the set of spots in M which contain I'. Then $M_{I'}$ is a model over I'.*

This model $M_{I'}$ is called the *reduced model* of M over I'.

Proof. Since $M_{I'}$ is a subset of a model (over I), we have only to show that $M_{I'}$ is the union of a finite number of affine models over I'. Therefore we may assume that M is an affine model. On the other hand, let A be the affine model defined by the affine ring $\mathfrak{S}$. Then $M_{I'}$ is contained in $J(A,M)$ and $M_{I'}=[J(A,M)\cap M]_{I'}$. Therefore we may assume further that the affine ring $\mathfrak{o}$ of M contains $\mathfrak{S}$. Let M' be the affine model defined by $I'[\mathfrak{o}]$ over I'. Since $I'[\mathfrak{o}]$ is a ring of quotients of $\mathfrak{o}$, M' is a subset of $M_{I'}$. Conversely, if a spot P is in $M_{I'}$, then P is a ring of quotients of $I'[\mathfrak{o}]$ and is in M'. Thus $M_{I'}=M'$ and the assertion is proved.

Now let $\mathfrak{p}$ be a prime ideal of I. Then the reduced model of M over $I_{\mathfrak{p}}$ is called the *local model* of M attached to the ground place $I_{\mathfrak{p}}$; it will be denoted by $M_{\mathfrak{p}}$.

Remark 1. If I is a field, then $\mathfrak{p}$ must be zero, and $M_{\mathfrak{p}}=M$.

Remark 2. If I is a semi-local ring, then any local model is again a model over I.

For, any ground place of I is an affine ring over I.

Remark 3. If I is a Dedekind domain with infinitely many prime ideals, then for any prime ideal $\mathfrak{p}$ of I, the local model $M_{\mathfrak{p}}$ is not a model over I, but is the intersection of infinitely many models over I.

Proof. Let $\boldsymbol{k}$ be the field of quotients of I. Assume that $M_{\mathfrak{p}}$ is a model over I. Let A' be an affine model over I which is contained in $M_{\mathfrak{p}}$ and let $\mathfrak{o}'$ be the affine ring of A'. Since $\boldsymbol{k}$ is an affine ring over $I_{\mathfrak{p}}$ and since $\mathfrak{o}'$ contains $I_{\mathfrak{p}}$, $\mathfrak{o} = \boldsymbol{k}[\mathfrak{o}']$ is an affine ring over I. Let A be the affine model over I defined by $\mathfrak{o}$. Then every spot in A has dimension at least 1, which is a contradiction by virtue of Proposition 3. Therefore $M_{\mathfrak{p}}$ is not a model over I. Now we prove the last assertion. For every maximal ideal $\mathfrak{q}$ of I other than $\mathfrak{p}$, we take an element a of $\mathfrak{q}$ which is not in $\mathfrak{p}$ and set $I((\mathfrak{q})) = I[1/a]$. Let $M(\mathfrak{q})$ be the reduced model of M over $I((\mathfrak{q}))$. Then each $M(\mathfrak{q})$ is a model over I and $M_{\mathfrak{p}}$ is the intersection of all $M(\mathfrak{q})$, which proves the last assertion.

9. Equivalence of the notions of models and of abstract varieties. Let V be a variety (in the sense of Weil [12]) defined over a field $\boldsymbol{k}$; we shall call such a variety an *affine variety*. Let (x) be a generic point of V over $\boldsymbol{k}$. Then the field $\boldsymbol{k}(x)$ is uniquely determined to within isomorphisms over $\boldsymbol{k}$ and is, by definition, a regular extension of $\boldsymbol{k}$. For a point P of V, we denote by $Q_V(P)$ the specialization ring of P in $\boldsymbol{k}(x)$ over $\boldsymbol{k}$. Then the set $S(V)$ of specialization rings $Q_V(P)$ $(P \varepsilon V)$ is the affine model defined by the affine ring $\boldsymbol{k}[x_1, \cdots, x_n]$ (where $(x) = (x_1, \cdots, x_n)$). Further, it is evident that for two points P and P' of V, P' is a specialization of P over $\boldsymbol{k}$ if and only if $Q_V(P')$ is a specialization of $Q_V(P)$. Conversely, if M is the affine model of a function field L over a ground field $\boldsymbol{k}$ defined by an affine ring $\boldsymbol{k}[x_1, \cdots, x_n]$ and if L is a regular extension of $\boldsymbol{k}$, and if V is the affine variety defined over $\boldsymbol{k}$ with generic point $(x_1, \cdots, x_n)$, then $S(V) = \{Q_V(P);\ P \varepsilon V\}$ coincides with M.

Next, let V and V' be birationally equivalent affine varieties and let T be a birational correspondence between V and V'. Assume that V, V' and T are defined over a field $\boldsymbol{k}$. Let (x) and (x') be corresponding generic points of V and V' under T. We can identify $\boldsymbol{k}(x)$ and $\boldsymbol{k}(x')$ by the correspondence T. Then it will be easily seen that the join of the affie models defined by $\boldsymbol{k}[(x)]$ and $\boldsymbol{k}[(x')]$ corresponds to T in the sense we stated above and that a point P in V corresponds biregularly to a point P' in V' if and only if $Q_V(P) = Q_{V'}(P')$, provided that P and P' are corresponding points under T. (Observe that if $Q_V(P) = Q_{V'}(P')$ for points P and P' in V and V' respectively, then there exists a point P'' in V' such that P'' is a generic specialization of P' over $\boldsymbol{k}$ and that P'' corresponds biregularly to P under T.)

Now let V be an abstract variety (in the sense of Weil [12]) defined over a field $\boldsymbol{k}$, with representatives V_i, frontiers $\mathfrak{F}_i$ and birational corre-

spondences T_{ij}. Let $(x^{(i)})$ be generic points of the varieties V_i which correspond to each other under the correspondences T_{ij}. Then we can identify the fields $\boldsymbol{k}(x^{(i)})$ with each other. Let L be the field obtained in this manner. Now let $P^{(i)}{}_1, \cdots, P^{(i)}{}_{n(i)}$ be points in $\mathfrak{F}_i$ such that every point of $\mathfrak{F}_i$ is a specialization of one of $P^{(i)}{}_1, \cdots, P^{(i)}{}_{n(i)}$ over $\boldsymbol{k}$; such points exist because $\mathfrak{F}_i$ is normally algebraic over $\boldsymbol{k}$ by definition. Then $S(\mathfrak{F}_i) = \{Q_{V_i}(P); P \varepsilon \mathfrak{F}_i\}$ is the union of the loci of the spots $Q_{V_i}(P^{(i)}{}_j)$ in the model $S(V_i)$ and is a closed set of $S(V_i)$. Therefore, $S(V_i - \mathfrak{F}_i) = \{Q_{V_i}(P); P \varepsilon V_i - \mathfrak{F}_i\}$ is a model of L by Theorem 9. Let P be a point of V. Then there exists one i such that P has a representative P_i in $V_i - \mathfrak{F}_i$. We call the specialization ring $Q_{V_i}(P_i)$ of P_i the specialization ring of P and denote it by $Q_V(P)$; even if P has another representative P_j in $V_j - \mathfrak{F}_j$, $Q_V(P)$ is unique because P_i and P_j correspond biregularly to each other. Now let $S(V)$ be the set of specialization rings of points in V. Then $S(V)$ is the union of all $S(V_i - \mathfrak{F}_i)$. Since $S(V_i - \mathfrak{F}_i)$ is a model, it is the union of a finite number of affine models, and the same is true of $S(V)$. To prove that $S(V)$ is a model, we have only to show that any two different spots in $S(V)$ do not correspond to each other. Assume that a spot $Q_V(P)$ corresponds to a spot $Q_V(P')$ $(P, P' \varepsilon V)$. Let P_i and P'_j be representatives of P and P' in $V_i - \mathfrak{F}_i$ and $V_j - \mathfrak{F}_j$ respectively. Then there exists a point P'' in $V_j - \mathfrak{F}_j$ which corresponds to P_i under T_{ij} such that P'' is a generic specialization of P'_j. Since P'' corresponds to P_i under T_{ij}, P'' is a representative of P, whence $Q_V(P) = Q_{V_j}(P'') \varepsilon S(V_j - \mathfrak{F}_j)$. Since $S(V_j - \mathfrak{F}_j)$ is a model, we have $Q_V(P) = Q_V(P')$ and $S(V)$ is a model.

Conversely, assume that M is a model of a function field L over a field $\boldsymbol{k}$ and assume that L is a regular extension of $\boldsymbol{k}$. Let $A_1, \cdots, A_n$ be affine models of L such that M is the union of them. Let $\mathfrak{o}_i = \boldsymbol{k}[x^{(i)}{}_1, \cdots, x^{(i)}{}_{m(i)}]$ be the affine ring of A_i. Let V_i be the affine variety of generic point $(x^{(i)}) = (x^{(i)}{}_1, \cdots, x^{(i)}{}_{m(i)})$ defined over $\boldsymbol{k}$ and let T_{ij} be the birational correspondence between V_i and V_j under which $(x^{(i)})$ and $(x^{(j)})$ correspond to each other. Then there exists an abstract variety V with representatives V_i, birational correspondences T_{ij} and empty frontiers. The variety V being defined in this manner, we see easily that $S(V)$, as defined above, coincides with M.

Remark. Let V be an abstract variety defined over a field $\boldsymbol{k}$. Then there exists an abstract variety V' which is everywhere in biregular correspondence with V and which has representatives V_i and correspondences T_{ij} such that the frontiers are empty.

Appendix. A proof of a result due to Krull.

LEMMA 1. *Let $\mathfrak{o}$ be a complete local integral domain of rank r. If $\mathfrak{p}_0 \subset \mathfrak{p}_1 \subset \cdots \subset \mathfrak{p}_s$ is a maximal chain of prime ideals in $\mathfrak{o}$ (that is, $\mathfrak{p}_0 = 0$, $\mathfrak{p}_s$ is maximal and each $\mathfrak{p}_i/\mathfrak{p}_{i-1}$ is of rank 1 for every $i = 1, \cdots, s$), then $r = s$.* (Cohen [2])

Proof. Let $\mathfrak{r}$ be an unramified regular local ring contained in $\mathfrak{o}$ such that $\mathfrak{o}$ is a finite $\mathfrak{r}$-module (Lemma 0.8). Since a regular local ring is a normal ring (Lemma 0.7) and since $\mathfrak{o}$ is an itegral extension of $\mathfrak{r}$, $\mathfrak{p}_1 \cap \mathfrak{r}$ is of rank 1 ([10, § 5]). Therefore $\mathfrak{p}_1 \cap \mathfrak{r}$ is a principal ideal (Lemma 0.9),[4] whence rank $(\mathfrak{r}/\mathfrak{p}_1 \cap \mathfrak{r})) = r - 1$ and rank $\mathfrak{o}/\mathfrak{p}_1 = r - 1$. Thus we prove our assertion easily by induction on r.

Now we prove

LEMMA 2. *Let $\mathfrak{r}$ be a regular local ring of rank n. If $\mathfrak{p}$ is a prime ideal of $\mathfrak{r}$, then* rank $\mathfrak{p}$ + co-rank $\mathfrak{p} = n$. (Krull [5])

Proof. Let $\mathfrak{r}^*$ be the completion of $\mathfrak{r}$. Then $\mathfrak{r}^*$ is also a regular local ring, and it is an integral domain (Lemma 0.7). Set rank $\mathfrak{p} = r$, co-rank $\mathfrak{p} = s$. Since $\mathfrak{r}^*/\mathfrak{p}\mathfrak{r}^*$ is the completion of $\mathfrak{r}/\mathfrak{p}$, rank $\mathfrak{r}^*/\mathfrak{p}\mathfrak{r}^* = s$ (Lemma 0.1), which shows that there exists a prime divisor $\mathfrak{p}^*$ of $\mathfrak{p}\mathfrak{r}^*$ such that co-rank $\mathfrak{p}^* = s$ and that there exists no prime divisor of $\mathfrak{p}\mathfrak{r}^*$ whose co-rank is greater than s. Therefore rank $\mathfrak{p}\mathfrak{r}^* = n - s$ by Lemma 1. On the other hand, let S be the complements of $\mathfrak{p}$ in $\mathfrak{r}$. Since $\mathfrak{r}^*/\mathfrak{p}\mathfrak{r}^*$ is the completion of $\mathfrak{r}/\mathfrak{p}$, every element of $\mathfrak{r}/\mathfrak{p}$ is not a zero-divisor in $\mathfrak{r}^*/\mathfrak{p}\mathfrak{r}^*$ (Corollary to Lemma 0.4), which shows that every element of S is not in any prime divisor of $\mathfrak{p}\mathfrak{r}^*$. Therefore rank $\mathfrak{p}\mathfrak{r}^* =$ rank $\mathfrak{p}\mathfrak{r}^*_S$. Let $a_1, \cdots, a_r$ be a system of parameters of $\mathfrak{r}_S = \mathfrak{r}_\mathfrak{p}$. Then $\mathfrak{p}\mathfrak{r}^*_S$ and $\sum a_i\mathfrak{r}^*_S$ have the same radical, whence rank $\mathfrak{p}\mathfrak{r}^*_S =$ rank $\sum a_i\mathfrak{r}^*_S$. Therefore rank $\mathfrak{p}\mathfrak{r}^*_S \leqq r$; it follows that rank $\mathfrak{p}\mathfrak{r}^* \leqq r$. Therefore $n - s \leqq r$. Since obviously $n \geqq r + s$, we see that $n - s = r$ and the assertion is proved.

MATHEMATICAL INSTITUTE,
KYOTO UNIVERSITY, JAPAN.

---

[4] If we want to avoid making use of Lemma 0.9, we may choose $\mathfrak{r}$ such that some element x of $\mathfrak{p}_1$ is in $\mathfrak{r}$ but not in the square of the maximal ideal of $\mathfrak{r}$. Then $\mathfrak{p}_1 \cap \mathfrak{r}$ is generated by x.

REFERENCES.

[1] C. Chevalley, "On the theory of local rings," *Annals of Mathematics*, vol. 44 (1943), pp. 690-708.

[2] I. S. Cohen, "On the structure and ideal theory of complete local rings," *Transactions of American Mathematical Society*, vol. 57 (1946), pp. 54-106.

[3] I. S. Cohen and A. Seidenberg, "Prime ideals and integral dependence," *Bulletin of American Mathematical Society*, vol. 52 (1946), pp. 252-261.

[4] W. Krull, "Idealtheorie," *Ergebnisse der Mathematik und ihrer Grenzgebiete*, vol. 1, No. 3 (1935).

[5] ———, "Dimensionstheorie in Stellenringen," *Journal für die reine und angewandte Mathematik*, vol. 179 (1938), pp. 204-226.

[6] T. Matusaka, "Some theorems on abelian varieties," *Natural Science Report of the Ochanomizu University*, vol. 4 (1953), pp. 22-35.

[7] M. Nagata, "Some remarks on local rings," *Nagoya Mathematical Journal*, vol. 6 (1953), pp. 53-58.

[8] ———, "Some remarks on local rings, II," *Memoirs of the College of Science, University of Kyoto*, ser. A, vol. 28, No. 2 (1954), pp. 109-120.

[9] ———, "Local rings" (I and II), *Sûgaku*, vol. 5 (1953-1954), pp. 104-114 and pp. 229-238 (in Japanese).

[10] ———, "Basic theorems on general commutative rings," forthcoming. (This paper appeared in *Memoirs of the College of Science, University of Kyoto*, ser. A, vol. 29, No. 1 (1955), pp. 59-77. (Added in October, 1955.))

[11] P. Samuel, "Algèbre locale," *Mémorial des Sciences Mathématiques*, No. 123 (1953).

[12] A. Weil, "Foundations of algebraic geometry," *American Mathematical Society Colloquium Publications* (1946).

[13] O. Zariski, "Some results in the arithmetic theory of algebraic varieties," *American Journal of Mathematics*, vol. 61 (1939), pp. 249-294.

[14] ———, "Foundations of general theory of birational correspondences," *Transactions of American Mathematical Society*, vol. 53 (1943), pp. 490-542.

[15] ———, "Analytical irreducibility of normal varieties," *Annals of Mathematics*, vol. 49 (1948), pp. 352-361.

[16] ———, "Sur la normalité analytique des variété normales," *Annales de L'Institut Fourier*, vol. 2 (1950), pp. 161-164.

[17] ———, "Theory and applications of holomorphic functions on algebraic variety over arbitrary ground fields," *Memoirs of the American Mathematical Society*, No. 5 (1951).

ON THE CURVATURES OF A SURFACE.*

By Aurel Wintner.

I. On the Mean Curvature.

1. Let S be a (piece of a) surface in the X-space, where $X=(x,y,z)$, and let $S\,\varepsilon\,C^n$ for a fixed $n>0$. By this is meant that a neighborhood of every point of S has *some* C^n-paramerization, that is, a parametrization of the form

$$(1)\qquad S\colon X=X(u,v),\qquad (u,v)\,\varepsilon\,D,$$

where D is an open, simply connected (u,v)-domain and $X(u,v)$ a vector function satisfying the conditions $X(u,v)\,\varepsilon\,C^n$ and $[X_u,X_v]\neq 0$ on D. Note that if $n<m$, then a C^n-parametrization of an $S\,\varepsilon\,C^m$ need not be a C^m-parametrization. If $S\,\varepsilon\,C^1$, there exists on S a continuous normal vector N. The latter is a function $N(u,v)\,\varepsilon\,C^{n-1}$ in terms of every C^k-parametrization (1) of an $S\,\varepsilon\,C^n$ not only if $k=n$ but also if $k=n-1$ (but not necessarily if $k=n-2$). In order that $S\,\varepsilon\,C^n$, where $n>1$, it is sufficient (and, of course, necessary) that S possesses some parametrization (1) satisfying

$$(2)\qquad X(u,v)\,\varepsilon\,C^{n-1} \text{ and } N(u,v)\,\varepsilon\,C^{n-1}, \text{ where } [X_u,X_v]\neq 0.$$

In fact, the sufficiency of (2) for *some* C^n-parametrization of S is a consequence of the (local) theorem on implicit functions; cf., e.g., [8], pp. 133-134, and the references given there.

If $a=a(u,v)$ denotes the matrix of the first, and $\beta=\beta(u,v)$ that of the second, fundamental matrix in terms of a C^2-parametrization (1) of an $S\,\varepsilon\,C^2$, then a is a (symmetric) positive definite matrix satisfying $a(u,v)\,\varepsilon\,C^1$ but nothing more than $\beta(u,v)\,\varepsilon\,C^0$ is in general true of the (symmetric) matrix β; so that the mean curvature $H=H(u,v)$ exists but is, in general, just continuous, since it is defined by

$$(3)\qquad H=\tfrac{1}{2}\operatorname{tr}(\beta a^{-1});$$

in view of

$$(4)\qquad K=\det(\beta a^{-1}),$$

* Received July 27, 1955.

this applies to the Gaussian curvature $K = K(u, v)$ also. As is well-known,

(5) $$K \leqq H^2,$$

where the sign of equality holds at a point (u, v) if and only if (u, v) is an umbilical point, that is, a point at which β becomes a scalar multiple of α (a "flat" point, with $K = 0 = H$, is here included as "umbilical").

The following considerations on H, considerations in which $S \varepsilon C^2$ or, if $S \varepsilon C^2$, the C^2-character of a given parametrization (1) of $S \varepsilon C^2$ will not be assumed, have various motivations. On the one hand, it can happen that no C^2-parametrization of a given $S \varepsilon C^2$ is available, since precisely the "natural" parametrization of $S \varepsilon C^2$, one given in geometrical terms, must fail to be a C^2-parametrization (cf. Section 8 below); so that the definition of (3) of H fails (in terms of what is available). On the other hand, the theory of minimum surface ($H \equiv 0$) speaks since Haar (cf. [5]) of surfaces $S \varepsilon C^1$ for which $S \varepsilon C^2$ is not assumed (even though proved, as a consequence of the theory). Such occurrences call for a definition of $H = H(u, v)$ which is more geometrical than (3) (concerning the geometrical, but erroneous, attempts of Minding and R. Sturm, cf. the comments of Stäckel [15]).

An approach of the desired type can be abstracted from the manner in which H appears when, in the classical instance of varying a double integral, the first variation,

(6) $$\delta \int_B\!\!\int |[X_u, X_v]| \, du dv,$$

of the area (of a piece of S) is transformed into a line integral.

2. Let (1) be a C^1-parametrization of an $S \varepsilon C^1$; so that $N(u, v) \varepsilon C^0$ holds for the unit vector

(7) $$N = [X_u, X_v]/g, \text{ where } g = |[X_u, X_v]| > 0$$

(this g is the positive square root of the determinant of

(8) $$|dX(u, v)|^2 = E(u, v) du^2 + 2F(u, v) du dv + G(u, v) dv^2,$$

the first fundamental form of S, with continuous E, F, G). By the existence of a continuous mean curvature on $S \varepsilon C^1$ should be meant that, for *some* continuous vector function $(\cdot \cdot \cdot)$ of (u, v),

(9) $$\int_J [N, X_u du + X_v dv] = \int_B\!\!\int (\cdot \cdot \cdot) \, du dv$$

holds as an identity in J and its interior $B=B(J)$. Here J denotes any positively oriented, piecewise smooth Jordan curve which, together with B, is contained in the fixed (u,v)-domain D on which the C^1-function (1) is given.

It is easy to realize that the existence of a continuous $(\cdot\cdot\cdot)$ satisfying the requirement (9) is not implied by the C^1-character of (1). If there exists a continuous $(\cdot\cdot\cdot)$, it is unique, since B is arbitrary. It also follows that $(\cdot\cdot\cdot)$ must be a vector function which is a (possibly vanishing) scalar multiple of $N(u,v)$, simply because the scalar product of $dX=X_u du+X_v du$ and N is 0; cf. (7). Hence $(\cdot\cdot\cdot)$ can be written as gN times $-2H$, where $H=H(u,v)$ is a unique continuous scalar, since $g=g(u,v)$ in (7) is positive and continuous. Let H be declared to be the mean curvature of S. The explicit form of (9) becomes

$$(10)\qquad \int_J [N, dx] = \int\int_B -2HgN\,du\,dv.$$

3. This transcribes into a (generalized) definition a fact pointed out by Weatherburn ([16], p. 255; cf. also [14]), who observed that (10) holds on every $S\,\varepsilon\,C^2$ if H is defined by (3). For reasons of covariance, (10) holds for C^1-parametrizations (1) of $S\,\varepsilon\,C^2$ also. Actually, it is clear from (7) that, for reasons of covariance, (10) holds in every C^1-parametrization of every $S\,\varepsilon\,C^1$ if it holds in one C^1-parametrization of that $S\,\varepsilon\,C^1$, and that H is invariant if a C^1-parametrization (1) is replaced by any other C^1-parametrization of $S\,\varepsilon\,C^1$ (provided that the local C^1-transformation, say

$$(11)\qquad u^*=u^*(u,v),\qquad v^*=v^*(u,v),$$

which represents this reparametrization is of positive Jacobian; if the Jacobian of (11) is negative, then N goes over into $-N$, hence H into $-H$, and so only H^2 is invariant).

If $S\,\varepsilon\,C^2$, and if (1) is a C^2-parametrization of S, then Gauss' definition of K (in terms of oriented areas of the spherical images of portions of S), when combined with the derivation formulae of Weingarten (cf., e.g., [2], p. 62), can be written in the form

$$(12)\qquad \int_J [N, dN] = \int\int_B -2KgN\,du\,dv.$$

This striking analogue of (10), also pointed out in Weatherburn's paper [16], can be considered as defining K in the same way as (10) defines H. But

(12), in contrast to (10), differentiates $N(u,v)$ and represents, therefore, a requirement which cannot be formulated if only $X(u,v)\,\varepsilon\,C^1$ is assumed. On the other hand, (12) holds whenever (1) is a C^1-parametrization of an $S\,\varepsilon\,C^2$. This follows for reasons of invariance, and since $N(u,v)\,\varepsilon\,C^1$ holds for every C^1-parametrization (1) of an $S\,\varepsilon\,C^2$ (conversely, the case $n=2$ of the assumptions (2) implies that $S\,\varepsilon\,C^n=C^2$).

A definition of the mean curvature which is more general than the definition of a *continuous* H (on an $S\,\varepsilon\,C^1$) above results as follows: With reference to any piecewise smooth J, put

$$\phi(B)=\int_J [N,dX], \tag{13}$$

where B is the interior of J. Suppose that the set-function (13) can be extended from the particular (u,v)-sets $B=B(J)$ to all Borel sets E, leading to an additive set-function $\phi(E)$, defined on the Borel field of (u,v)-sets E contained in D. Then, if $\phi(E)$ is absolutely continuous, the analogue of the relation (10) leads to an L-integrable $H(u,v)$.

4. It will now be proved that an $S\,\varepsilon\,C^1$ possessing a continuous H need not be an $S\,\varepsilon\,C^2$; in other words, that the definition, (10), of Section 2 is actually more general than the classical definition, (3). In fact, the example $S\,\varepsilon\,C^1$ to be given will be such that $\beta(u,v)$ will in no sense exist at a point, say $(u,v)=(0,0)$, although there will exist a continuous H. Incidentally, this particular S will be such as to possess a C^1-parametrization in which $\alpha(u,v)\,\varepsilon\,C^1$ (as though (1) were a C^2-parametrization, which (1) cannot be, since $S\,\varepsilon\,C^2$ fails to hold).

First, if an $S\,\varepsilon\,C^2$ is given in a Cartesian parametrization, say as $z=z(x,y)$ (so that $u=x$, $v=y$), then $z(x,y)\,\varepsilon\,C^2$ on an (x,y)-domain, D, and (3) reduces to

$$(1+q^2)r-2pqs+(1+p^2)t=2Hh^3 \tag{14}$$

where $p=z_x,\cdots,t=z_{yy}$ and

$$h=(1+p^2+q^2)^{\frac{1}{2}}. \tag{15}$$

Next, if $[F]$ denotes the Lagrangian derivative of an $F=F(x,y,z,p,q)$, then, formally, the identical vanishing of (6) is equivalent to the case $F=h$ of $[F]=0$, whereas Laplace's equation, $r+t=0$, belongs to $F=p^2+q^2$. Since $[F+G]=[F]+[G]$, and since $[G]=-f(x,y)$ when $G=$

$G(x,y,z,p,q)$ is of the form $G = zf(x,y)$, the formal analogy between (14) and Poisson's equation

(16) $$r + t = f(x,y)$$

indicates that, in order to obtain an $S \varepsilon C^1$ of the desired type, a construction of Lichtenstein [10], which refers to $F = p^2 + q^2 - 2zf(x,y)$, can be adapted.

Lichtenstein's construction is based on the results of Petrini [12] on logarithmic potentials. The following construction will have the same basis, in a form similar to that used in [8], pp. 134-135.

5. Let D be the unit circle $0 \leqq x^2 + y^2 < 1$, and let $w(x,y)$ be $(-2\pi)^{-1}$ times the logarithmic potential of a density $f(x,y)$ which is uniformly continuous on D. According to Petrini, $w(x,y) = w_f(x,y)$ need not possess second derivatives $r, \cdots$ (so that (16) need not be meaningful for $z = w$), if f is suitably chosen on D, an example being

(17) $$f(x,y) = (\cos\theta)^2/\log\rho \qquad (x = \rho\cos\theta, y = \rho\sin\theta)$$

(if $0 < \rho$, with $f(0,0) = 0$ at $\rho = 0$); cf. [12], p. 138. But $z = w$ must always satisfy the "integrated form" of (16), which is

(18) $$\int_J (pdy - qdx) = \iint_B fdxdy$$

(cf. [10], pp. 99-100), where J and $B = B(J)$ are meant in the same sense as in (9). It is also known that, since $f(x,y)$ is uniformly continuous, $p = w_x$ and $q = w_y$ will not only exist but be such as to satisfy a uniform Hölder condition of any index $\lambda < 1$; in particular, of some index $\lambda > \frac{1}{2}$. But the analyticity of (17) for $\rho \neq 0$ implies that $w_f(x,y)$ is analytic, hence of class C^2, at every point $(x,y) \neq (0,0)$ of D. On the other hand, since $w(x,y)$ is the logarithmic potential of a continuous density, it is readily seen that $w_x(x,y) = O(|\log r|)$ and $w_y(x,y) = O(|\log r|)$ as $r = (x^2+y^2)^{\frac{1}{2}} \to 0$ (in this connection, cf. [18], pp. 736-737). Hence, if $p = w_x$ and $q = w_y$ are such as to satisfy

(19) $$p(0,0) = 0 \text{ and } q(0,0) = 0,$$

then

(20) $$p^2(x,y) \varepsilon C^1, \quad p(x,y)q(x,y) \varepsilon C^1, \quad q^2(x,y) \varepsilon C^1 \text{ on } D\colon 0 \leqq x^2 + y^2 < 1$$

(even though $p(x,y) \varepsilon C^1$, $q(x,y) \varepsilon C^1$ cannot hold, since $w(x,y) \varepsilon C^2$ is false). Finally, (19) becomes satisfied, and both (20) and the formulation (18) of

(16) remain valid, if $w(x,y)$ in the preceding deduction is replaced by $z(x,y)+cx+dy$, where $c=-w_x(0,0)$, $d=-w_y(0,0)$.

Accordingly, $S\colon z=z(x,y)$ is of class C^1 but not of class C^2 (even though the coefficients of (8), where $(u,v)=(x,y)$, are functions of class C^1, since

$$E=1+p^2, \qquad F=pq, \qquad G=1+q^2, \tag{21}$$

and since (20) holds). It remains to be shown that this S has a continuous $H=H(x,y)$.

6. A Pfaffian

$$a(x,y)\,dx+b(x,y)\,dy \tag{22}$$

is called regular (on D) if $a(x,y)$, $b(x,y)$ are continuous and such as to satisfy the condition

$$\int_J (adx+bdy)=\int\int_B fdxdy \tag{23}$$

for *some* continuous $f(x,y)$, where $J, B=B(J)$ have the same meaning as in (9). It is clear from Green's theorem that (22) must be regular if $a(x,y)\in C^1$ and $b(x,y)\in C^1$. It is also known that if (22) is regular (for some reason), then $\mu(x,y)$ times the Pfaffian (22) must also be regular whenever

$$\mu(x,y)\in C^1, \tag{24}$$

since (23) and (24) imply that

$$\int_J \mu\cdot(adx+bdy)=\int\int_B (\mu f-a\mu_y+b\mu_x)\,dxdy. \tag{25}$$

This is a theorem of E. Cartan (which unfortunately escaped us in [6], where, without a reference to [3], pp. 69-70, it is stated (p. 761) and, essentially with E. Cartan's proof, is proved as a lemma).

Let $\mu=1/h$. Then (15) and (20) show that (24) is satisfied. Hence, if (23) is identified with (18), the case $\mu=1/h$ of (25) is applicable and leads to

$$\int_J h^{-1}\cdot(z_xdy-z_ydx)=\int\int_B (\cdot\ \cdot\ \cdot)\,dxdy, \tag{26}$$

where $(\cdot\ \cdot\ \cdot)$ is a certain continuous function of (x,y). On the other hand, since the C^1-parametrization (1), where $X=(x,y,z)$, is now given in the form

$$S\colon z=z(x,y), \tag{27}$$

it is clear that the vector requirement (10) for a continuous H reduces to the scalar condition

$$(28)\qquad \int_J (1+z_x^2+z_y^2)^{-\frac{1}{2}}(z_x dy - z_y dx) = \int\int_B 2\dot{H}\,dxdy.$$

Since the existence of a continuous H satisfying (28) follows from (26) and (15), this proves the existence of an S which has the properties announced at the beginning of Section 4.

7. A matrix function

$$(29)\qquad \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

of (u,v) is called regular (on a (u,v)-domain D) if both Pfaffians

$$a(u,v)\,du + b(u,v)\,dv, \qquad c(u,v)\,du + d(u,v)\,dv$$

are regular. Since this will be the case if (though not only if) the function (29) is of class C^1, the matrix, $\alpha(u,v)$, of the first fundamental form (8) of a C^2-parametrization (1) of an $S \varepsilon C^2$ is a regular matrix. With regard to the matrix, $\beta(u,v)$, of the second fundamental form

$$(30)\qquad L(u,v)\,du^2 + 2M(u,v)\,dudv + N(u,v)\,dv^2,$$

the situation is as follows:

The classical formulation of the Mainardi-Codazzi equations is of the form

$$(31)\qquad L_v - M_u = (\cdot\cdot\cdot), \qquad M_v - N_u = (\cdot\cdot\cdot)$$

and assumes a C^3-parametrization (1) for an $S \varepsilon C^3$. Under this assumption, $\alpha(u,v) \varepsilon C^2$ and $\beta(u,v) \varepsilon C^1$. It follows therefore from Green's identity that (31) is equivalent to the pair of relations

$$(32)\qquad \begin{aligned} \int_J (Ldu + Mdv) &= \int\int_B (\cdot\cdot\cdot)\,dudv, \\ \int_J (Mdu + Ndv) &= \int\int_B (\cdot\cdot\cdot)\,dudv \end{aligned}$$

(as identities in J and its interior B), where the expressions $(\cdot\cdot\cdot)$ are the same as the respective expressions $(\cdot\cdot\cdot)$ in (31). Both of these functions $(\cdot\cdot\cdot)$ of (u,v) are bilinear forms in

$$(33)\qquad (L, M, N) \text{ and } (E_u, E_v, F_u, F_v, G_u, G_v),$$

with coefficients which are rational functions of the coefficients of (8). There is no point in writing here down this classical pair of expressions $(\cdot\ \cdot\ \cdot)$, the more so as, in the more general formulation of (32) which will be given under (II) below, $(\cdot\ \cdot\ \cdot)$ *could not be written down* (in fact, the second of the two sets (33) will then become undefined). The preceding assumption, $S \varepsilon C^3$, is relaxed to $S \varepsilon C^2$ in (I) and (II) below.

(I) Let (1) be a C^2-parametrization of an $S \varepsilon C^2$. Then all $3+6$ functions (33) exist and are continuous, since $\beta(u,v) \varepsilon C^0$, $a(u,v) \varepsilon C^1$. Moreover, if the two classical expressions $(\cdot\ \cdot\ \cdot)$ are formed from the functions (33) in the same way as in the traditional case $S \varepsilon C^3$, it can be shown (cf. [6], pp. 760-766) that the Mainardi-Codazzi equations are valid in the form (32). Consequently, $\beta(u,v)$ (and not only $a(u,v)$; cf. above) is a regular matrix if (1) is a C^2-parametrization of an $S \varepsilon C^2$.

(II) It follows that (32) holds in terms of every C^1-parametrization (1) of an $S \varepsilon C^2$, provided that the (continuous) elements L, M, N of the (symmetric) matrix function β are defined, not directly (which, in terms of the C^1-parametrization of $S \varepsilon C^2$, would not be possible), but by covariance, and that the validity of the Mainardi-Codazzi equations is interpreted to mean the existence of *certain* continuous functions $(\cdot\ \cdot\ \cdot)$ for which the pair of relations (32) holds as an identity in J and its interior B. In terms of a C^2-parametrization (1) of $S \varepsilon C^2$, the second of the sets (33) enters into $(\cdot\ \cdot\ \cdot)$ in the form of the Christoffel coefficients Γ^i_{jk} of $a(u,v) \varepsilon C^1$. But Γ^i_{jk} is not a tensor; its transformation rule involves the *second* derivatives of a C^1-transformation (11) of non-vanishing Jacobian, and these second derivatives will not exist if (11) represents the passage from a C^2-parametrization (1) to a C^1-parametrization $X = X(u^*; v^*)$ of $S \varepsilon C^2$. On the other hand, the coefficients of (30) form a contravariant tensor under C^1-transformation (11) (of positive Jacobian) and are, therefore, defined in terms of C^1-parametrization of $S \varepsilon C^2$ also. Since the regularity of any matrix (29) is preserved under C^1-transformations (11) of non-vanishing Jacobian (when (29) is thought of as a tensor which is contravaniant in both indices), the proof is complete.

8. It seems to be quite artificial to consider, as in (II) or (12) and (10), parametrizations (1) of an $S \varepsilon C^2$ which are just C^1-parametrizations. Actually, such a step is often unavoidable, since it can be imposed by the *geometrical* structure of a problem. A convincing instance of this situation presents itself in the theory of ruled surfaces, a theory which (as it turns out,

partly for this reason) is in a notoriously poor shape from the point of view of analysis.

Let

(34) a ruled surface R of class C^2

be defined as follows: R is an $S \varepsilon C^2$ possessing a C^1-parametrization (1) in which the parameter lines $u = \text{const.}$ correspond to segments $\Lambda = \Lambda(u)$ of straight lines in the X-space, where $X = (x, y, z)$. In other words, (34) means that R is a surface S which, besides possessing *some* C^2-parametrization (1), has a C^1-parametrization of the form

$$R: X(u,v) = A(u)v + B(u), \tag{35}$$

where $A(u)$, $B(u)$ are vector functions possessing continuous first derivatives $A'(u)$, $B'(u)$ (on a certain interval $u_0 < u < u^0$).

It would be out of place to replace this definition of (34) by the requirement that $R \varepsilon C^2$ should have a C^2-parametrization of the form (35). For suppose that an R possessing a C^2-parametrization happens to be a "developable" (a "torse"), in the sense that the plane tangent to R at (u, v) does not vary along $\Lambda(u)$. Then it is readily seen from (35), where $A(u) \varepsilon C^2$ and $B(u) \varepsilon C^2$ by assumption, that the normal vector (7) is a function $N(u)$ of class C^2. Since the case $n = 3$ of (2) implies that $S \varepsilon C^3$, it follows that R, instead of being just an $S \varepsilon C^2$, must be an $S \varepsilon C^3$. For a certain converse (even with C^2, C^3 replaced by C^1, C^2, respectively), cf. [8], pp. 133-134. In what follows, neither $N(u, v) = N(u)$ nor $R \varepsilon C^3$, but only (34), will be assumed.

Since (35) is just a C^1-parametrization, the coefficients of the first fundamental form (8) follow from (35) as continuous functions, whereas the second fundamental form cannot be formed in this parametrization. But the (continuous) second derivatives $X_{vv}, X_{uv} = X_{uv}$ of (35) exist (though X_{uu} need not), and so it is easy to realize that the last two coefficients, M and N, of (30) can be defined not only by covariance (as in (II), Section 7), but also by their standard definition. According to the latter, M, N are the scalar products of $[X_u, X_v]$ and X_{uv}/g, X_{vv}/g, respectively, where $g = (EG - F^2)^{\frac{1}{2}} > 0$ (cf., e. g., [2], p. 52). In view of (35), this leads to

$$M = \det(A, A', B')/g, \quad N = 0, \tag{36}$$

where $g = |[X_u, X_v]|$ and $' = d/du$. Hence, in order to obtain (30) in terms of the C^1-parametrization (35) of $R \varepsilon C^2$, only $L = L(u, v)$ remains to be determined. But it turns out that L is given by

$$L = 2g^2H/G + 2F\det(A, A', B')/(gG), \tag{37}$$

where the coefficients of (8) are supplied by the first derivatives of (35) and $H = H(u, v)$ is the mean curvature. The latter is supplied by (10), since (10) is valid in terms of the C^1-parametrization (35) of $R \varepsilon C^2$ also.

In order to verify (37), note that, since the coefficient matrices, α and β, of (8) and (30) are contravariant tensors, (3) holds in terms of C^1-parametrizations of every $S \varepsilon C^2$. But (37) follows if (36) is inserted in the explicit form of (3).

9. In Section 8, reference was made to a paradoxical situation which originates from the conclusion $(2) \rightarrow (S \varepsilon C^n)$. In what follows, the same conclusion will lead to certain, geometrically quite unexpected, implications dealing with the degree of smoothness concerning both types of classical developables, say $S = V = V(\Gamma)$ and $S = W = W(\Gamma)$, which are the envelopes (V) of the normal planes and the envelopes (W) of the rectifying planes of a space curve Γ.

The results to be obtained can be considered as counterparts of the results obtained, in [19], p. 368, and [20], p. 251, respectively, for the parallel surfaces (Steiner) and the evolute surfaces (Monge) of a surface. In all four cases, the result is to the effect that, subject to a rank condition on the Jacobian matrix involved, the "generated" surfaces are smoother than one would expect from their geometrical definitions.

Let Γ: $X = X(s)$, where $X = (x, y, z)$, be a (sufficiently short piece of a) curve of class C^3 possessing a non-vanishing second derivative $X''(s)$ with respect to the arc length s. Then Γ has a positive curvature $\kappa(s) \varepsilon C^1$ and a torsion $\tau(s) \varepsilon C^0$, and the unit vectors $U_1 = X'$, $U_2 = X''/\kappa$, $U_3 = [U_1, U_2]$ satisfy Frenet's equations

$$(38) \qquad U'_1 = \kappa U_2, \qquad U'_2 = -\kappa U_1 + \tau U_3, \qquad U'_3 = -\tau U_2$$

(so that $U_1(s) \varepsilon C^2$ but $U_i(s) \varepsilon C^1$ if $i = 2$ or $i = 3$). Conversely, if any positive function $\kappa(s) \varepsilon C^1$ and any real-valued function $\tau(s) \varepsilon C^0$ are given on an s-interval, then an application of the existence and uniqueness theorem of linear systems of ordinary differential equations supplies a unique $\Gamma \varepsilon C^3$ satisfying (38).

10. It follows that if only $\Gamma \varepsilon C^3$ (with $\kappa > 0$) is assumed, then the function

$$(39) \qquad X(s, t) = X(s) + U_2(s)/\kappa(s) + tU_3(s),$$

where t is a linear parameter and $X = X(s)$ is the parametrization of Γ in

terms of the arc length, need not have a second derivative $X_{ss}(s,t)$. But it turns out that the locus S: $X = X(s,t)$, which, as is well-known (cf., e.g., [2], pp. 43-44), is the envelope $S = V = V(\Gamma)$ defined in Section 9, is an $S \varepsilon C^2$, provided that the curve, say $a(s,t) = 0$, on which

$$(40) \qquad [X_s, X_t] \neq 0$$

fails to hold is excluded from S. The explicit representation of a is

$$(41) \qquad a = a(s,t) = t\tau(s) + \kappa'(s)/\kappa^2(s)$$

(if $a(s,t) = 0$ is inserted from (41) into (39), it follows that that singular curve on S is the path, if any, of the centres of the osculating spheres of Γ; cf., e.g., [2], p. 36, and, if $\Gamma \varepsilon C^n$ for some $n \geqq 4$, the theorem italicized on p. 246 of [20], where $n = 4$).

Accordingly, the "natural" parametrization, (39), of $S = V(\Gamma)$, being just a C^1-parametrization, disguises the fact that, under the restriction (40), there must exist some C^2-parametrization $X = X(u;v)$. The proof will be such as to show a corresponding result, $S \varepsilon C^{n-1}$ (instead of just $S \varepsilon C^{n-2}$) when $\Gamma \varepsilon C^n$ in $S = V(\Gamma)$, holds for every $n \geqq 3$.

Suppose that $\Gamma \varepsilon C^3$ and $(\kappa > 0)$. Then (38) is applicable. But if (39) is differentiated with respect to s and t and then use is made of (38), where $U_1(s) = X'(s)$, it is seen that

$$(42) \qquad [X_s, X_t] = -aU_1,$$

where a is the scalar defined by (41). It follows from (42) that, on the one hand, $S = V(\Gamma)$ has a (unit) normal vector $N = N(s,t)$ unless $a = a(s,t)$ vanishes and that, on the other hand, $N = \pm U_1$ if $a \neq 0$, that is, if (40) is satisfied (it also follows that $N(s,t) = N(s)$, but this will not be used). Since $U_1(s) \varepsilon C^2$, it follows that the normal vector N is a function of class C^2, and therefore of class C^1, in a C^1-parametrization, (39), of $S = V(\Gamma)$. Hence $S \varepsilon C^2$, as claimed.

11. If the envelope $V(\Gamma)$ is replaced by the envelope $W(\Gamma)$ (cf. Section 9), then (39) becomes replaced by

$$(43) \qquad X(s,t) = X(s) + t\{\tau(s)U_1(s) + \kappa(s)U_3(s)\}$$

(cf., e.g., [2], pp. 45-46). If only $\Gamma \varepsilon C^3$ (and, as always, $\kappa(s) > 0$) is assumed, then, since the torsion $\tau(s)$ can be any continuous function (cf. the remark made at the end of Section 9), the continuous function (43) need not be differentiable.

Suppose therefore that $\Gamma \varepsilon C^4$. Then $\tau(s) \varepsilon C^1$ (and $\kappa(s) \varepsilon C^2$). But since $\tau(s) \varepsilon C^2$ need not hold, (43) will not in general be a function of class C^2. Accordingly, the "natural" parametrization, (43), of $S = W(\Gamma)$, where $\Gamma \varepsilon C^4$, is just a C^1-parametrization, provided that the inequality (40) (which, if

$$\beta = \beta(s,t) = \kappa(s) + t\{\tau'(s)\kappa(s) - \tau(s)\kappa'(s)\}, \tag{44}$$

turns out to be equivalent to $\beta \neq 0$) is satisfied. Nevertheless, $W(\Gamma) \varepsilon C^2$ (as long as $\beta \neq 0$).

In fact, if (38) is used in the same way as in Section 10, it follows from (43) that what now correspond to (42) and (41) are

$$[X_s, X_t] = -\beta U_2 \tag{45}$$

and (44), respectively, Hence, if the "path" of the points (s, t) at which (44) vanishes is excluded from the surface $W(\Gamma)$, then $W(\Gamma) \varepsilon C^2$, where $\Gamma \varepsilon C^4$, follows from (45) in the same way as $V(\Gamma) \varepsilon C^2$, where $\Gamma \varepsilon C^3$, followed from (42) in Section 10. It is also seen that, as long as $\beta(s, t)$ does not vanish, $n = 4$ can be replaced by any $n \geqq 4$ in the assertion $W(\Gamma) \varepsilon C^{n-2}$, if $\Gamma \varepsilon C^n$ is the assumption.

12. Let $S \varepsilon C^*$ mean that $S \varepsilon C^1$ and that S possesses a continuous mean curvature (in the sense of Section 2). Thus $S \varepsilon C^2$ implies that $S \varepsilon C^*$ but the converse is not true (Sections 4-5). This situation leads to various unanswered questions of which only three, (α), (β) and (γ) below, will be mentioned.

(α) If an $S \varepsilon C^*$ is a convex surface (in the sense that its tangent planes, which exist, since $S \varepsilon C^1$ by virtue of $S \varepsilon C^*$, support S), must S be of bounded curvature K in the sense of A. D. Alexandroff [1], Chap. XI and pp. 491-514 (and must this K satisfy (5))? It is natural to raise this question since $K \geqq 0$ holds formally but the proof of (5), when based on (3)-(4), asumes that $S \varepsilon C^2$. In this connection, cf. also (+?) in [22], p. 848.

(β) If $S \varepsilon C^*$, must S possess a C^1-parametrization (1) in terms of which the first fundamental form (8) becomes isothermic (i. e., $E = G$, $F = 0$)? This question has some interest in view of [4], where H is replaced by K. An affirmative answer would be surprising, since H, in contrast to K, has little to do with the metric (8).

(γ) Does $S \varepsilon C^1$ imply $S \varepsilon C^2$ if S possesses continuous curvatures H, K and is free of "umbilical points," points at which the sign of equality holds in (5)? The answer to those variants of this question in which the indices of differentiability are raised is known to be affirmative (cf. [8], p. 128). It is understood that, in the present case, K must be thought of as defined in terms of the metric (8) of $S \varepsilon C^*$; for instance, by assuming that S is a C^1-embedding of a regular C^1-metric (in this connection, cf. also (iv?) in [22], pp. 846-848). Incidentally, it is now not clear that the sign of inequality cannot reverse itself in (5); cf. question (β). Correspondingly, (γ) has an analogue in the direction of (β).

Other problems are those in the large, such as the uniqueness statement of Christoffel and the corresponding existence question for convex surfaces $S \varepsilon C^*$. The latter question concerns the existence of a closed, strictly convex $S \varepsilon C^*$ for which H is assigned as a (positive, continued) function of the normal. The method of construction which in Part II below will be applied to another embedding problem (Weyl) supplies a counterexample only if H is allowed to have an infinity. A similar situation prevails if H is replaced by K (Minkowski).

*Remark.** Suppose that an $S \varepsilon C^2$ has a C^2-representation (27) over a domain D of the (x, y)-plane and let J be a rectifiable Jordan curve which, along with its interior $B = B(J)$, is contained in D. It was recently observed by E. Heinz (*Mathematische Annalen*, vol. 129 (1955), pp. 451-454) that

$$1/r \geqq \min_{B+J} |H|, \tag{46}$$

if J is (or surrounds) a circle of radius r.

Since the proof of (46) depends on (28), which is (10) in the case (27), and since (10) is paralleled by (12), it is natural to raise the question concerning a counterpart of (46) in which H is replaced by K. It will turn out that the answer to this question is

$$|I|/(2\pi r^2) \geqq \min_{B+J} |K|, \tag{47}$$

if $|I|$ denotes the length of I, where I is the image (on the unit sphere $|N| = 1$) of J under the Gaussian mapping $N = N(u, v)$ of S: $X = X(u, v)$.

Actually, both (46) and (47) can be generalized to the case in which

---

* Added November 19, 1955.

J is any rectifiable Jordan curve. In fact, if $|J|$ denotes the length of J, and $|B|$ the area of the (x,y)-domain B surrounded by J, then

$$|J|/|B| \geqq \min_{B+J} 2|H| \tag{48}$$

and

$$|I|/|B| \geqq \min_{B+J} 2|K|. \tag{49}$$

Clearly, (48) reduces to (46), and (49) to (47), if J is a circle of radius r. The proof of (48) (which is between the lines of Heinz's proof of (46)) and of (49) proceeds as follows:

Since $|[Y,Z]| \leqq |Y||Z|$ and $|N|=1$, the absolute values of the line integrals occurring in (10) and (12) are majorized by

$$\int_J |dX| = |J| \text{ and } \int_J |dN| = |I|, \tag{50}$$

respectively. On the other hand, since (1) is given in the form (27), the relation (10) simplifies to (28), and (12) to a relation similar to (28) (with $2H$ replaced by $2K$ in the $dxdy$-integral). Accordingly, (10), (12) and the majorants (50) of the line integrals lead to

$$\left|\int_B 2Hdxdy\right| \leqq |J|, \qquad \int_B 2Kdxdy \leqq |I|. \tag{51}$$

Clearly, (48) and (49) follow from (51).

Heinz points out (loc. cit., Satz 2) that (46) can be combined with (5) if S satisfies the convexity assumption $K>0$. Under the same assumption, (48), too, can be combined with (5) and leads to

$$(\tfrac{1}{2}|J|/|B|)^{2} \geqq \min_{B+J} K > 0 \tag{52}$$

(whereas the classical isoperimetric inequality, being the relation $|J|/|B|^{\frac{1}{2}} \geqq (4\pi)^{\frac{1}{2}}$, cannot be combined with (48) or (52), at least not directly). But (52) is of a type quite different from (49) and, in contrast to (52), neither (49) nor (48) asumes that S is convex.

II. Embedding and Gaussian Curvature.

1. In general terms, Weyl's embedding problem can be formulated as follows: On a two-dimensional manifold $\mathfrak{S}$ which is of the topological type of a sphere, there is assigned a metric μ which (in sense to be specified)

has a positive curvature. The problem consists in finding in the Euclidean (x, y, z)-space a surface Σ which is topologically equivalent to Θ and which realizes μ on Σ by virtue of the topological correspondence between Θ and Σ.

From the point of view of differentiability assumptions, there is today a considerable gap between what is known to be true and what is known to be false concerning the existence of a Σ (cf. [11] and [7], respectively). Corresondingly, the assumptions of smoothness on μ under which the existence of a Σ of appropriate smoothness is assured today, appear to be quite restictive from the point of view of *local* embeddings (in this regard, cf. [22]; in the existence proof given in [11], the strictness of the assumptions placed on μ is originated, as in [17], by the necessity of appealing to Weyl's inequality (cf. [23]) in the treatment of the *non-local* problem).

No such analytical restrictions on μ, and no corresponding complications in the proof, arise in A. D. Alexandroff's approach to the problem (cf. [1], Chap. VII; cf. Chap. XI). This approach, like Minkowski's treatment of his embedding problem, obtains Σ as a limit of closed convex polyhedrons $\Pi_1, \Pi_2, \cdots$. Correspondingly, since the strength of the limit process $\Pi_i \rightarrow \Sigma$ is not under control from the point of view of differentiability (as a matter of fact, Π_i itself is not a differentiable surface), no assertion of smoothness can result for Σ.

The following comments have the purpose of exhibiting the intrinsic necessity of not claiming any smoothness for Σ (at least when the curvature $K > 0$ of μ is allowed to be ∞ at a point of Θ; except for this point, μ will be a regular analytic Riemannian metric).

2. Let D be a domain, say the circle $D = D_a$: $u^2 + v^2 < a^2$, in a parameter plane (u, v), and let there be given on D a metric ds^2, that is, a quadratic form (8) (but no function X), with coefficients E, F, G which are continuous on D and satisfy the condition $g > 0$, where $g = (EG - F^2)^{\frac{1}{2}} > 0$. Then ds^2 will be called a continuous metric (on D). By a C^1-embedding S of ds^2 in the X-space, where $X = (x, y, z)$, is meant a surface S: $X = X(u, v)$, where $X(u, v)$ is a vector function of class C^1 satisfying (8). If $X(u, v)$ is a function of class C^n, then the surface S is called a C^n-embedding of ds^2.

Let $B = B_a$ denote the domain which results if the point $(u, v) = (0, 0)$ is removed from the circle $D = D_a$, so that B_a: $0 < u^2 + v^2 < a^2$, and let ds^2 be a metric which is continuous on D_a: $u^2 + v^2 < a^2$, with coefficients E, F, G which are of class C^2 (or, for that matter, regular analytic) on B_a and possess a curvature $K = K(u, v)$ which tends to ∞ as $u^2 + v^2 \rightarrow 0$ (so that $K > 0$ on B_a, if a is small enough). It will be shown that, roughly speaking,

such a continuous metric ds^2 on D_a can be chosen in such a way that for no positive $c(<a)$ will the metric ds^2 on D_c possess any *convex* C^1-embedding S (even though there exists on B_c an analytic $K(u,v)$ having a positive lower bound, and even though E, F, G are continuous on D_c or on the closure of B_c).

Since $K>$ const. >0, the italicized proviso, that of restricting the C^1-embeddings S to be convex, seems to be redundant (the more so as the metric is regular analytic on B_c). But as matters seem to stand today, this conclusion is valid only if the surfaces S admitted are assumed to be C^2-embeddings, rather than just C^1-embeddings (in this regard, cf. (iv?) and (+?) in [22], p. 848.* As a matter of fact, the proof to be given below will assume that the C^1-embedding S of the continuous metric ds^2 is a C^2-embedding near all those points (u,v) at which the functions E, F, G are not of class C^2 (and these functions will be regular analytic except at a single point, a point at which they will remain continuous). All such embeddings S turn out to be convex (as a matter of fact, strictly convex) by virtue of $K>0$; cf. the proof in Section *4* below.

3. Let $E=1$ and $F=0$, hence $g=G^{\frac{1}{2}}$; so that

$$ds^2=du^2+g^2dv^2, \tag{1}$$

where $g=g(u,v)>0$. Then, at those points (u,v) at which g has continuous second derivatives, $K=K(u,v)$ is given by

$$g_{uu}+Kg=0 \tag{2}$$

(Jacobi). Define on D_a a continuous metric (1) by placing $g(0,0)=1$ at D_a-B_a and

$$g(u,v)=1+(\cos\phi)^2/\log r, \text{ where } u=r\cos\phi, v=r\sin\phi \tag{3}$$

$(r>0)$; so that $g(u,v)$ is positive and regular analytic on B_a (if $a>0$ is small enough). Incidentally, cf. (*3*) with (17) above.

Substitution of (*3*) into (*2*) shows that $K(u,v)\sim-2/(r^2\log r)$ as $r\to\infty$. This implies that $K>0$ on B_a (if $a>0$ is small enough) and that

* As will, however, be shown elsewhere, a treatment of questions (+?)-(—?) of [22], p. 848, and of the related questions, pointed out above, can be based on the results of A. D. Alexandroff [1], p. 51, concerning convex metrics.

the integral of $K(u,v)$ over B_a is ∞. Since $g(u,v)\to 1$ as $(u,v)\to(0,0)$, this is equivalent to

$$\iint\limits_{B_a} Kg\,du\,dv=\infty \qquad (K>0, g>0). \tag{4}$$

It will be concluded from (4) that, no matter how small $a>0$ be chosen, there cannot belong to (1) on D_a any surface $S\colon X=X(u,v)$ of class C^1 satisfying $ds=|dX|$, provided that, corresponding to (but, as far as present knowledge seems to go, perhaps not implied by) the fact that the coefficient g of (1) is regular analytic, and K is positive, on B_a (with $K=\infty$ at D_a-B_a), that portion, say S_0, of the C^1-surface S which corresponds to B_a is assumed to be of class C^2 (hence, strictly convex).

4. Suppose the contrary. Then, since S is a surface of class C^1, it can be assumed to be in the form $S\colon z=z(x,y)$, where x,y,z are Cartesian coordinates, $(u,v)=(0,0)$ corresponds to $(x,y,z)=(0,0,0)$, the function $z(x,y)$ is of class C^1 in an (x,y)-neighborhood, say V, of the point $(x,y)=(0,0)$ and, if Q denotes the latter point, the (x,y)-plane is tangent to S at Q (i.e., $z_x=0$ and $z_y=0$ at Q). In terms of the notation S_0, introduced at the end of Section 3, the surface S_0 results if the origin is removed from the surface S. Since S_0 is supposed to be of class C^2, the function $z(x,y)$ has continuous second derivatives on the (x,y)-domain $V-Q$.

According to Satz IV of Schilt [13], p. 257, the origin is the only point of S on the (x,y)-plane (if V or $a>0$ is small enough). In fact, $z(x,y)$ is of class C^2, hence the normal vector to S_0 is of class C^1, on $V-Q$, and the proof, given in [13], is such that the "gradient," considered on p. 244, need not be assumed to exist at the critical point. Accordingly, it can be assumed that $z(x,y)>0$ on $V-Q$, since $z(x,y)=0$ at Q.

Also the consideration of Schilt on pp. 250-251 of [13] (§9-§10) remain valid, since, for the proofs given loc. cit., only the existence and the continuity, but not the differentiability, of the normal to S is needed at Q. Hence, Satz V of Schilt [13], p. 257, is applicable.

This means that, if $N=N(x,y)$ denotes the (oriented) unit normal at the point (x,y) of S (where the boundary point $(0,0)$ of $S-S_0$ is included), then $N=N(x,y)$ represents a one-to-one continuous mapping of a neighborhood V of Q on the unit sphere. Let T denote the spherical image of V or S, and let $|T|$ be the area of R. Then, since T is a schlicht image of V, and since $K>0$ on $V-Q$,

$$|T|=\lim_{\epsilon\to 0}\iint\limits_{B_a-B_\epsilon} Kg\,du\,dv$$

(Gauss). It follows therefore from (*4*) that $|T| = \infty$. But $|T| = \infty$ contains a contradiction. In fact, since T is a schlicht piece of the unit sphere, $|T| < 4\pi$.

5. The preceding results is of a local nature, since it deals only with a sufficiently small D_a. It is, however, clear that the case (*3*) of (*1*) can be extended from D_a to a closed, orientable (u, v)-manifold in such a way as to be of class C^∞ and of positive $K(u, v)$ on the closure of $\Sigma - D_a$, and the passage from the class C^∞ to the class of regular analyticity (except for the center of V_a) also offers no difficulty. What thus results is the situation announced in Section *2*.

ADDENDUM.* Let there be given on the abstract sphere Θ a positive definite metric form ds^2, suppose that the coefficients E, F, F of ds^2 are functions of class C^1 in suitable local parameters (u, v) on Θ and that ds^2 has at every point of Θ a continuous curvature K, finally that $K > 0$ throughout. The general existence statement in Weyl's paper [17] is that, under these assumptions, the ds^2 on Θ can be realized on a (strictly) convex, closed surface Σ of class C^2 in the X-space, where $X = (x, y, z)$. This existence statement of Weyl will be referred to as (*).

Weyl was aware that he did not fully succeed in proving (*) (cf. [11], where further references are given), and it was observed in [7] that certain variants of (*), variants the truth of which one would expect if (*) is true, are certainly false. It will now be shown that (*) itself is false.

For a real $z = z(x, y)$, and on a sufficiently small neighborhood of $(x, y) = (0, 0)$, consider the implicit equation

$$(5) \qquad z = \tfrac{1}{2}(x^2 f + y^2/f), \text{ where } f = f(z) = 2 + \sin\log(1/\log z);$$

so that $z(x, y) > 0$ unless $(x, y) = (0, 0)$, and $z(0, 0) = 0$. Let D be the circle $0 \leqq r < a$, and D_0 the punctured circle $0 < r < a$, where $r = (x^2 + y^2)^{\frac{1}{2}}$. According to A. D. Alexandroff [1], pp. 446-447 (where $\log w$ must be interpreted as $\log - w$ if $w < 0$), the implicit relation (*5*) defines over D a differentiable convex cap S: $z = z(x, y)$ which is such that, although the plane curves which represent the normal sections of S at the point $D - D_0$ fail to have curvatures,

$$(6) \qquad S \text{ possesses a continuous } K > 0 \text{ on } D$$

(hence, at the point $D - D_0$ also). Clearly, S is analytic (hence $S \varepsilon C^2$) on D_0, but what was mentioned before (*6*) (and, in fact, (*5*) as it stands) shows

* Added November 19, 1955.

that $S \varepsilon C^2$ cannot hold on D. In what follows, it will be granted that, as claimed by Alexandroff, (5) defines over D a convex cap S satisfying (6).

Let every O refer to $(x,y) \to (0,0)$, i.e., to $r \to 0$. Then it is clear from (5) that $z(x,y) = O(r^2)$. Since S is convex, this implies that the derivatives $p = z_x$, $q = z_y$, which are analytic on D_0, are of the form $O(r)$ near $D - D_0$. By a standard property of functions which are derivatives of a continuous function, this implies that p and q exist ($=0$) and are continuous at $D - D_0$. Hence $S \varepsilon C^1$ on D.

Since the Gaussian parameters (u,v) are (x,y) in the C^1-parametrization $z = z(x,y)$ of S, the (continuous) coefficients of ds^2 on D are given by (21) above. On the other hand, the argument which led to (20) is applicable, since p and q are analytic, hence of class C^1, on D_0 (though not on D) and continuous and of the form $O(r)$ at $D - D_0$ (actually, only $p = o(r^{\frac{1}{2}})$ and $q = o(r^{\frac{1}{2}})$ are needed). Hence, E, F, G are of class C^1 on D.

The situation can be summarized as follows: On D, both $S \varepsilon C^1$ and (6) hold, but $S \varepsilon C^2$ does not hold (except on D_0), although S possesses on D a C^1-parametrization in which the coefficients of the ds^2 on S are functions of class C^1. In order to conclude from this that the assertion (*) is false, it is sufficient to repeat the argument applied on p. 487 in [7].

THE JOHNS HOPKINS UNIVERSITY.

REFERENCES.

[1] A. D. Alexandrow, *Die innere Geometrie der konvexen Flächen*, Berlin, 1955.

[2] L. Bieberbach, *Differentialgeometrie*, Teubner's Leitfäden no. 31 (1932).

[3] E. Cartan, *Leçons sur les invariants intégraux*, 1922.

[4] S.-S. Chern, P. Hartman and A. Wintner, "On isothermic coordinates," *Commentarii Mathematici Helvetici*, vol. 28 (1954), pp. 301-309.

[5] A. Haar, "Zur Variationsrechnung," *Abhandlungen aus dem Mathematischen Seminar der Hamburgischen Universität*, vol. 8 (1931), pp. 1-27.

[6] P. Hartman and A. Wintner, "On the fundamental equations of differential geometry," *American Journal of Mathematics*, vol. 72 (1950), pp. 757-774.

[7] ——— and A. Wintner, "On pieces of convex surfaces. Appendix," *ibid.*, vol. 75 (1953), pp. 485-487.

[8] ——— and A. Wintner, "On the curvatures of a surface," *ibid.*, vol. 75 (1953), pp. 127-141.

[9] L. Lichtenstein, "Über das Verschwinden der ersten Variation bei zweidimensionalen Variationsproblemen," *Mathematische Annalen*, vol. 69 (1910), pp. 514-516.

[10] ———, "Über einige Integrabilitätsbedingungen zweigliedriger Differentialausdrücke mit einer Anwendung auf den Cauchyschen Integralsatz," *Sitzungsberichte der Berliner Mathematischen Gesellschaft*, vol. 9 (1910), pp. 84-100.

[11] L. Nirenberg, "The Weyl and Minkowski problems in differential geometry in the large," *Communications on Pure and Applied Mathematics* (New York), vol. 6 (1953), pp. 337-394.

[12] H. Petrini, "Les dérivées premières et secondes du potential logarithmique," *Journal de Mathématique*, ser. 6, vol. 5 (1909), pp. 127-223.

[13] H. Schilt, "Über die isolierten Nullstellen der Flächenkrümmung und einige Verbiegbarkeitssätze," *Compositio Mathematica*, vol. 5 (1937), pp. 239-283.

[14] E. Scholz, "Flächentheoretische Integralsätze," *Schriften des Mathematischen Seminars der Universität Berlin*, vol. 1 (1933), pp. 161-185.

[15] P. Stäckel, "Grenzübergänge der Krümmungslehre," *Jahresbericht der deutschen Mathematiker-Vereinigung*, vol. 27 (1918), pp. 142-157.

[16] C. E. Weatherburn, "On differential invariants in geometry of surfaces, with some applications to Mathematical Physics," *Quarterly Journal of Mathematics*, vol. 50 (1925), pp. 230-269.

[17] H. Weyl, "Über die Bestimmung einer geschlossenen konvexen Fläche durch ihr Linienelement," *Vierteljahrsschrift der Naturforschenden Gesellschaft in Zürich*, vol. 61 (1915), pp. 40-72.

[18] A. Wintner, "On the Hölder restrictions in the theory of partial differential equations," *American Journal of Mathematics*, vol. 72 (1950), pp. 731-738.

[19] ———, "On parallel surfaces," *ibid.*, vol. 74 (1952), pp. 365-376.

[20] ———, "On the infinitesimal geometry of curves," *ibid.*, vol. 75 (1953), pp. 241-259.

[21] ———, "On the local rôle of the theory of the logarithmic potential in differential geometry," *ibid.*, vol. 75 (1953), pp. 679-690.

[22] ———, "On the local embedding problems in the differential geometry of surfaces," *American Journal of Mathematics*, vol. 77 (1955), pp. 845-852.

[23] ———, "On Weyl's identity in the differential geometry of surfaces," to appear in the *Annali di Matematica*.

A VARIATIONAL METHOD IN THE THEORY OF HARMONIC INTEGRALS, II.* [1]

By Charles B. Morrey, Jr.

---

1. Introduction. In this part, the variational method introduced in part I [9] is applied to the study of boundary value problems for exterior differential forms on a compact Riemannian manifold $\mathfrak{M}$ with boundary $\mathfrak{B}$. The manifold is not assumed to be orientable and parallel theories are developed for even and odd forms.

We shall use the results of part I extensively and shall refer to it frequently; the words part I in such a reference will stand for that paper [9]. We retain the notations of that part except for one change used only in Sections 2 and 3 and introduce new notations as required.

In Section 2, we discuss the behavior of the G-quasi-potentials and G-potentials defined in Section 3 of part I on the part $x^n = 0$ of the boundary of a hemisphere. The behavior of certain more general quasi-potentials and potentials, satisfying a "natural boundary condition" on $x^n = 0$ is also studied. In Section 3, systems of equations in integrated form like those in Section 4 of part I are studied, particularly with reference to the differentiability properties of the solutions along $x^n = 0$. Certain approximation and boundedness theorems are also proved. The results of these two sections form the analytic basis for our results concerning the differentiability of the solutions at the boundary.

In Section 4 important preliminary material is presented: Riemannian manifolds with boundary of class $C_\mu{}^k$, etc., are defined, certain results about $\mathfrak{P}_2$ forms are carried over from Section 5 of part I, the tangential and normal parts of the boundary values of forms are defined, a number of important lemmas are proved, and the Gaffney inequality [5] and theorem are proved for the closed linear subspaces $\mathfrak{P}_2{}^+$ of $\mathfrak{P}_2$-forms ω for which $n\omega = 0$ and $\mathfrak{P}_2{}^-$ of $\mathfrak{P}_2$-forms ω for which $t\omega = 0$.

The Gaffney theorem for the spaces $\mathfrak{P}_2{}^+$ and $\mathfrak{P}_2{}^-$ makes it possible to

---

* Received May 24, 1955.

[1] This research was partially supported by the United States Air Force through the Office of Scientific Research of the Air Research and Development Command.

carry over verbatim the analysis of Section 6 of part I to each of the spaces $\mathfrak{P}_2^+$ and $\mathfrak{P}_2^-$. The differentiability theory of Section 3, together with the approximation device used in the proof of Theorem 7.1 of part I, is used to establish complete results concerning the differentiability properties of the "plus and minus potentials" and their derivatives. These results are then used to establish an orthogonal decomposition theorem $\mathfrak{L}_2 = \mathfrak{H} \oplus \mathfrak{C} \oplus \mathfrak{D}$ similar to that of Kodaira [6] given in part I in which $\mathfrak{H}$ consists of all harmonic fields in $\mathfrak{L}_2$, and $\mathfrak{C}$ and $\mathfrak{D}$ consist, respectively, of all $\mathfrak{L}_2$-forms of the form $\delta\alpha$ for α in $\mathfrak{P}_2^+$ and $d\beta$ for β in $\mathfrak{P}_2^-$. The differentiability of the projections of a given form on these spaces is discussed completely and several other interesting related theorems are proved. In this case, the manifold $\mathfrak{H}$ has infinite dimensionality, but the Friedrichs inequality [4] $D(\omega) \geqq \lambda \| \omega \|^2$ is proved to hold for all ω in $\mathfrak{C} \oplus \mathfrak{D}$. This leads directly to the existence of a certain overall potential Ω in $\mathfrak{C} \oplus \mathfrak{D}$ of forms ω in $\mathfrak{C} \oplus \mathfrak{D}$ which turns out to be related to the positive and negative potentials Ω^+ and Ω^- of ω; in fact

$$(1.1) \qquad d\Omega^+ = d\Omega, \qquad \delta\Omega^- = \delta\Omega.$$

In Section 6, we begin by deducing the results in Theorems 2, 3, and 4 of the paper [3] of Duff and Spencer very quickly from the Theorems of Section 5. Next the Dirichlet problem (tK and nK given) for harmonic *forms* (Theorem 1 of Duff and Spencer [3]), as distinct from harmonic fields, is proved including the result of Spencer [11] concerning the uniqueness of the solutions in the analytic case. The potentials Ω^+, Ω^-, and Ω of Section 5 satisfy the boundary conditions

$$(1.2) \qquad n\Omega^+ = nd\Omega^+ = 0; \quad t\Omega^- = t\delta\Omega^- = 0; \quad nd\Omega = t\delta\Omega = 0.$$

These potentials are used to obtain the results of Connor [1] concerning boundary value problems for harmonic forms K. He solves the problems (i) nK and ndK given, (ii) tK and $t\delta K$ given, and (iii) ndK and $t\delta K$ given. We solve these, roughly speaking, by showing first that there is some form ω satisfying the given boundary conditions and then defining K as ω minus the relevant potential of $\Delta\omega$. This procedure is carried through under very general conditions and complete results concerning differentiability are presented. We conclude with a remark concerning a recent extension due to Spencer [10] of the Dirichlet problem to bounded manifolds.

As was pointed out in the introduction to part I, K. O. Friedrichs had obtained some results for the case of manifolds without boundary but had not published them. He has been working almost concurrently on the case

of manifolds with boundary. All of his results are included in his paper [4]. After a detailed discussion with him of our and his results and methods, we have concluded that there are sufficient differences to warrant publication of both of our papers.

2. Potentials on hemispheres. In this section we study the G-quasi-potentials and G-potentials defined in part I [9], definition 3.4 for the case that G is a hemisphere G_R where G_R is the part of $B(0,R)$ for which $x^n < 0$. We also study certain unrestricted $G(x_0,R)$-quasi-potentials and potentials which we define below in Definition 2.2. In this and the next section, we retain the notations of Sections 3 and 4 of part I unless otherwise specified. For convenience, we denote by σ_R the part of $B(0,R)$ for which $x^n = 0$ and S_R^- the part of $\partial B(0,R)$ for which $x^n \leqq 0$. We also define $G(x,r)$ as the part of $B(x,r)$ for which $x^n < 0$ and $S^-(x,r)$ as the part of $S(x,r)$ for which $x^n \leqq 0$ and define $\Gamma(x,r)$ as the union of $G(x,r)$, its reflection $G^+(x,r)$ in $x^n = 0$, and the interior of the $(n-1)$-sphere $\partial G(x,r) \cap \partial G^+(x,r)$.

THEOREM 2.1. (i) *If* $u \in \mathfrak{P}_2$ *on* $G(x_0,R)$ *and vanishes on* $S^-(x_0,R)$, *then* $d_0[u, G(x_0,R)] \leqq C_1 R d_1[u, G(x_0,R)]$, $C_1 = 2^{\frac{1}{2}}$.

(ii) *The totality of such functions forms a closed linear manifold in* $\mathfrak{P}_2$.

Proof. (i) follows from the integration of the inequality

$$(2.1) \quad \int |u(x^n, x'_n)|^2 dx'_n = \int |u(x^n, x'_n) - u(X^n, x'_n)|^2 dx'_n$$
$$\leqq (x^n - x_0^n + R) D_2[u, G(x_0,R)],$$
$$X^n = x_0^n - (R^2 - |x'_n - x'_{0n}|^2)^{\frac{1}{2}}$$

with respect to x^n. (ii) follows from Theorem 2.12, part I.

Definition 2.1. We define the space $\mathfrak{P}^*_2$ on $G(x_0,R)$ to consist of all u in $\mathfrak{P}_2$ which vanish on $S^-(x_0,R)$ with inner product given by

$$((u,v))^* = \int_{G(x_0,R)} u_{x^\alpha} v_{x^\alpha} dx$$

and norm $\| u \|^* = ((u,u))^{\frac{1}{2}}$.

Remark. It is clear that $\mathfrak{P}_{20} \subset \mathfrak{P}^*_2$.

The proof of Theorem 3.3 of part I with the obvious modifications yields the following theorem:

THEOREM 2.2. *Suppose* $e=(e_1,\cdots,e_n)$ *and* f *are in* $\mathfrak{L}_2$ *on* $G(x_0,R)$. *Then there are unique solutions* u^* *and* v^* *in* $\mathfrak{P}^*_2$ *on* $G(x_0,R)$ *such that*

$$(2.2)\qquad \int_{G(x_0,R)} w_{x^\alpha}(u^*_{x^\alpha}+e_\alpha)\,dx=0,\ \int_{G(x_0,R)}(w_{x^\alpha}v^*_{x^\alpha}+wf)\,dx=0$$

for all w *in* $\mathfrak{P}^*_2$ *on* $G(x_0,R)$ *and we have*

$$\| u\|^*\leqq d_0[e,G(x_0,R)],\qquad \| v\|^*\leqq C_1Rd_0[f,G(x_0,R)].$$

Definition 2.2. The solution u^* above is called the $G(x_0,R)$*-*-quasi-potential of* e and the solution v^* is called the $G(x_0,R)$*-*-potential of* f.

The following lemma of Soboleff is proved exactly as it was in part I, Lemma 3.2:

LEMMA 2.1. *Suppose* $u,\nabla u,\cdots,\nabla^p u$ *are all of class* $\mathfrak{P}_2$ *on* $G(x_0,R)$ $p\geqq[n/2]$. *Then*

$$| u(x)|\leqq a_{n*}^{-\frac{1}{2}}R^{-n/2}\{\sum_{j=0}^{p}d_j[u,G(x_0,R)](2R)^j/j!$$
$$+d_{p+1}[u,G(x_0,R)]2\cdot(2R)^{p+1}/p!\};$$
$$x\,\varepsilon\, G(x_0,R),\qquad a_{n*}=\tfrac{1}{2}m[B(0,1)].$$

Remark. It is clear that the definitions and theorems above extend immediately to vector functions u, v, w, etc. This is true of the following theorems and definitions also.

The proof of Lemma 3.6, part I, carries over with only very minor changes to yield a proof of Lemma 2.2 below; in fact the writer originally proved the lemma for this case (see [7], pp. 130-132).

Definition 2.3. For points $x\,\varepsilon\, G(x_0,R)$, δ_x denotes the distance of x from $S^-(x_0,R)$.

LEMMA 2.2. *Suppose* $u\,\varepsilon\,\mathfrak{P}_2$ *on* $G(x_0,R)$ *and suppose*

$$d_1[u-l_{yr},G(y,r)]\leqq L(r/\delta_y)^{\rho+\mu},\quad 0\leqq r\leqq\delta_y,\quad 0<\mu<1,\quad \rho=n/2,$$

for each $y\,\varepsilon\, G(x_0,R)$ *where* l_{yr} *is the linear function of Lemma* 3.5, *part* I, *for* u *on* $G(y,r)$. *Then* u *is of class* C^1_μ *on* $G(x_0,R)$ *and there is a constant* $C_2=C_2(n,\mu)$ *such that*

$$|\nabla u(\xi)-\nabla u(x)|\leqq C_2L\delta_x^{-\rho-\mu}\cdot|\xi-x|^\mu,$$
$$0\leqq|\xi-x|\leqq\delta_x/2,\ \xi,x\,\varepsilon\, G(x_0,R).$$

Definition 2.4. We define the *spaces* $\mathfrak{L}_{2\lambda}$ and $C_\mu{}^0$ of functions in $\mathfrak{L}_2$ on $G(x_0, R)$ and the respective norms $|e|_\lambda$ and $|e|_\mu{}^0$ exactly as they were defined (for functions in $\mathfrak{L}_2$ on B_R) in Definition 3.5 of part I, δ_x having its significance above. We define the *spaces* $\mathfrak{P}_{20\lambda}$ and $C_{0\mu}{}^1$, and $\mathfrak{P}^*{}_{2\lambda}$ and $C^*{}_\mu{}^1$, and the corresponding norms just as the first two were defined in Definition 3.5 of part I with the obvious changes, the first two being subsets of $\mathfrak{P}_{20}$ and the latter two being subsets of $\mathfrak{P}^*{}_2$ on $G(x_0, R)$.

The writer found it convenient to replace the spaces $\mathfrak{A}_{RK}{}^0$ and $\mathfrak{A}_{0RK}{}^1$ of Section 3, part I, by other spaces which we define below in Definition 2.6 which definition necessitates the following:

Definition 2.5. By $'\nabla^p e$, $p = 1, 2, \cdots$, we mean the set of functions defined by $'\nabla^0 e = e$, $'\nabla^p e = \{e_{x\alpha\cdots x\gamma}\}$, $\alpha, \cdots, \gamma \leqq n-1$. We define the non-negative quantities $d^*{}_p(e, G)$ and $d^{**}{}_p(u, G)$ by

$$d^*{}_0(e, G) = d_0(e, G),\ [d^*{}_p(e, G)]^2 = \int_G [|'\nabla^p e|^2 + |'\nabla^{p-1} e_{xn}|^2]\,dx, \qquad p \geqq 1,$$

$$d^{**}{}_0(u, G) = d_1(u, G),\ [d^{**}{}_p(e, G)]^2 = \int_G [|'\nabla^p \nabla u|^2 + |'\nabla^{p-1}\nabla u_{xn}|^2]\,dx, \qquad p \geqq 1.$$

Definition 2.6. We define the *space* $\mathfrak{B}_{RK}{}^0$ to consist of all e (or f) which are in $\mathfrak{L}_2$ on G_R with all the $'\nabla e^p$ in $\mathfrak{P}_2$ on each G_r with $r < R$ with norm defined by

$$|e|_{RK}{}^0 = \sup[(2p)!K^p]^{-\frac{1}{2}} \cdot (R-r)^p d^*{}_p(e, G_r), \qquad p = 0, 1, 2, \cdots, 0 \leqq r < R.$$

We define the *spaces* $\mathfrak{B}^*{}_{RK}{}^1$ and $\mathfrak{B}_{RK}{}^1$ as those subsets of $\mathfrak{P}^*{}_2$ and $\mathfrak{P}_{20}$, respectively for which all the $'\nabla u^p$ and $'\nabla u_{xn}{}^p \in \mathfrak{P}_2$ on each G_r with $r < R$, the norm being defined by

$$\| u \|^*{}_{RK}{}^1,\ \| u \|_{0RK}{}^1 = \sup[(2p)!K^p]^{-\frac{1}{2}} \cdot (R-r)^p d^{**}{}_p(u, G_r), \qquad 0 \leqq r < R, p = 0, 1, 2, \cdots.$$

LEMMA 2.3. *If* $u \in \mathfrak{B}^*{}_{RK}{}^1$ *or to* $\mathfrak{B}_{0RK}{}^1$, *then all the* $'\nabla u^p$ *and* $'\nabla u^p{}_{xn}$ *are continuous in* x *on* G_R *and analytic in* x'_n *for each* x^n, $-R < x^n \leqq 0$.

Proof. Suppose $-R < a < b \leqq 0$, $r^2 + a^2 < R^2$. Let v be any one of the functions above. Since v is of class $\mathfrak{P}_2$ on any $G_{r'}$ with $r' < R$, we have

$$\int_{\sigma_r} |\bar v(b, x'_n) - \bar v(a, x'_n)|^2\, dx'_n < (b-a)\int_a^b \Big[\int_{\sigma_r} |\bar v_{xn}(x^n, x'_n)|^2\, dx'_n\Big]\,dx^n,$$

so that $\int_{\sigma_r} |\bar{v}(x^n, x'_n)|^2 dx'_n$, $r < R$, is *uniformly bounded* for $a \leqq x^n \leqq 0$ if $a^2 + r^2 < R^2$. Then Soboleff's lemma 2.1 in $(n-1)$ dimensions implies *uniform* (independent of x^n) bounds for all the $'\nabla^p u$ and $'\nabla^p u_{x^n}$ on any G_r with $r < R$ which insure their analyticity in the variables x'_n for each x^n. Since each is also absolutely continuous in x^n for almost all x'_n, the continuity in x follows from the uniform continuity (with respect to x^n) in the variables x'_n.

The following symmetry lemma is useful in the proof of our main Theorem 2.4:

LEMMA 2.4. *Suppose that e and $f \in \mathfrak{L}_2$ on $G(x_0, R)$, suppose that $u_0 = Q_{R0}(e)$, $v_0 = P_{R0}(f)$, $u^* = Q^*_R(e)$, $v^* = P^*_R(f)$, and suppose that U_0, V_0, U^*, V^*, E_0, F_0, E^*, and F^* are defined on $\Gamma(x_0, R)$ to be equal on $G(x_0, R)$ to the corresponding small lettered functions and to be defined on $G^+(x_0, R)$ by*

$$\begin{aligned} &U_0(x^n, x'_n) = -u_0(-x^n, x'_n), \quad U^*(x^n, x'_n) = u^*(-x^n, x'_n), \\ (2.3)\qquad &E_{0n}(x^n, x'_n) = e_n(-x^n, x'_n), \quad E^*_n(x^n, x'_n) = -e_n(-x^n, x'_n), \\ &E_{0\alpha}(x^n, x'_n) = -e_\alpha(-x^n, x'_n), \quad E^*_\alpha(x^n, x'_n) = e_\alpha(-x^n, x'_n), \\ &\qquad\qquad\qquad\qquad\qquad\qquad \alpha = 1, \cdots, n-1, \end{aligned}$$

with formulas for V_0 and F_0 like those for U_0 and formulas for V^* and F^* like those for U^*. Then $U_0 = Q_R(E_0)$, $V_0 = P_R(E_0)$, $U^* = Q_R(E^*)$, $V^* = P_R(F^*)$. Here Q_{0R}, P_{0R}, Q^*_R, and P^*_R refer to $G(x_0, R)$ and Q_R and P_R to $\Gamma(x_0, R)$.

Proof. All of these are proved in a similar way so we shall prove only the first. We note that ∇U_0 and ∇U^* are related to ∇u_0 and ∇u^* as E_0 and E^* are to e and e, respectively. Let w be any function in $\mathfrak{P}_{20}$ on $\Gamma(x_0, R)$; obviously (since $u_0 = v_0 = 0$ on $x^n = 0$), U_0, V_0, U^*, $V^* \in \mathfrak{P}_{20}$ on $\Gamma(x_0, R)$. We write $w = w_1 + w_2$ where w_1 is odd in x^n like U_0 and w_2 is even in x^n like U^*. Then $\nabla U_0 \cdot \nabla w_1$ and $\nabla w_1 \cdot E_0$ are even in x^n while $\nabla w_2 \cdot \nabla U_0$ and $\nabla w_2 \cdot E_0$ are odd; clearly $w_1 = 0$ on $\partial G(x_0, R)$. Clearly, then,

$$\int_{\Gamma(x_0,R)} w_{x^\alpha}(U_{0x^\alpha} + E_{0\alpha})\,dx = 2\int_{G(x_0,R)} w_{1x^\alpha}(u_{0x^\alpha} + e_\alpha)\,dx = 0.$$

We now list a few additional properties of functions in the various spaces; the analog of (ii) should have been included in Theorem 3.5 of part I.

THEOREM 2.3. (i) *There is a constant $C_3(n,\mu)$ such that*

$$(2.4) \qquad |e(x)| \leqq C_3 \cdot |e|_\mu^0 \delta_x^{-\rho}, \qquad x \in G(x_0, R),$$

for any $e \in C_\mu^0$ on $G(x_0, R)$.

(ii) *There is a constant $C_4(n,\mu)$ such that if $e \in C_\mu^0$ on $G(x_0,R)$, then $e \in \mathfrak{L}_{2\lambda}$ for any $\lambda \leqq \rho$ on $G(x_0,R)$ with $|e|_\lambda \leqq C_4 |e|_\mu^0$.*

(iii) *There is a constant $C_5(n,\mu)$ such that if $u \in C_{0\mu}^1$ or to $C^*{}_\mu^1$ on $G(x_0, r_0)$, then u and $\nabla u \in C_\mu^0$ there and*

$$|\nabla u|_\mu^0 \leqq \|u\|_\mu^1, \quad |u|_\mu^0 \leqq C_5 \cdot R \cdot \|u\|_\mu^1, \quad \|u\|_\mu^1 = \|u\|_{0\mu}^1 \text{ or } \|u\|^*{}_\mu^1.$$

(iv) *If $u \in \mathfrak{B}_{0RK}^1$ or $\mathfrak{B}^*{}_{RK}^1$, then u and $\nabla u \in \mathfrak{B}_{RK}^0$ and*

$$|\nabla u|_{RK}^0 \leqq \|u\|'_{RK}, \quad |u|_{RK}^0 \leqq [C_1 + (2K)^{-\frac{1}{2}}]R \cdot \|u\|'_{RK},$$
$$\|u\|'_{RK} = \|u\|_{0RK}^1 \text{ or } \|u\|^*{}_{RK}^1.$$

The proofs of (i), (iii), and (iv) are very similar to the corresponding parts of Theorem 3.5, part I. (ii) follows easily by squaring (2.4), integrating over $G(y,r)$, and using the finiteness of $|e| \leqq |e|_\mu^0$.

THEOREM 2.4. *The transformations Q_{R0} and P_{R0} ($Q^*{}_R$ and $P^*{}_R$) are bounded linear operators from $\mathfrak{L}_{2\lambda}$ to $\mathfrak{P}_{20\lambda}(\mathfrak{P}^*{}_{2\lambda})$ for $0 < \lambda < \rho$, from C_μ^0 to $C_{0\mu}^1(C^*{}_\mu^1)$ for $0 < \mu < 1$, and from $\mathfrak{B}_{RK}^0$ to $\mathfrak{B}_{0RK}^1(\mathfrak{B}^*{}_{RK}^1)$ for any $K > e$. The transformation $P_R(P^*{}_R)$ is also a bounded linear operator from $\mathfrak{L}_{2,\rho-1+\mu}$ to $C_{0\mu}^1(C^*{}_\mu^1)$ for $0 < \mu < 1$. There are constants C_6, C_7, C_8, C_9, $C^*{}_6$, $C^*{}_7$, $C^*{}_8$, and $C^*{}_9$ with $C_k = C_k(n,\mu)$ and $C^*{}_k = C^*{}_k(n,\mu)$, $k = 6, 7, 8$, and $C_9 = C_9(K,n)$, $C^*{}_9 = C^*{}_9(K,n)$ such that*

$$\|P_{R0}(f)\|_{0\lambda} \leqq C_6 R |f|_\lambda, \qquad \|P^*{}_R(f)\|^*{}_\lambda \leqq C^*{}_6 R |f|_\lambda,$$

$$\|P_{R0}(f)\|_{0\mu}^1 \leqq C_7 R |f|_\mu^0, \qquad \|P^*{}_R(f)\|^*{}_\mu^1 \leqq C^*{}_7 R |f|_\mu^0,$$

$$\|P_{R0}(f)\|_{0\mu}^1 \leqq C_8 R |f|_{\rho-1+\mu}, \qquad \|P^*{}_R(f)\|^*{}_\mu^1 \leqq C^*{}_8 R |f|_{\rho-1+\mu},$$

$$\|P_{R0}(f)\|_{0RK}^1 \leqq C_9 R |f|_{RK}^0, \qquad \|P^*{}_R(f)\|^*{}_{RK}^1 \leqq C^*{}_9 R |f|_{RK}^0.$$

Proof. We use the notation of Lemma 2.4. We see that the first results for Q_{R0} and $Q^*{}_R$ and the first three results for P_{R0} and $P^*{}_R$ follow from that lemma, Theorems 2.2 and 2.3, and Theorems 3.5 and 3.6 of part I. Using also Lemma 3.3 and the very last part of the proof of Theorem 3.6, both of part I, we see that

$$\nabla V_0 = Q_R(G_0) + H_0, \ \ \nabla V^* = Q_R(G^*) + H^*, \ \ G_{0\gamma\alpha} = \delta_{\gamma\alpha} F_0, \ \ G^*_{\gamma\alpha} = \delta_{\gamma\alpha} F^*.$$

$$L_2(H_0, B_R), L_2(H^*, B_R) \leqq Z_1^2(n) R^2 L_2(f, G_R).$$

We may therefore deduce the last statements for P_{R0} and P^*_R from the symmetry lemma (note the symmetry properties of G_0), Theorem 2.3, and the last $(\mathfrak{B}_{RK}{}^0)$ results for both Q_{R0} and Q^*_R.

In order to prove the last results for Q_{R0} and Q^*_R, we begin by applying Lemma 3.3 of part I *repeatedly with* $\gamma \leqq n-1$ to the functions U_0 and U^*; it is clear from the proof of that lemma that this procedure is valid. Exactly as in the proof of (iii) in Theorem 3.6 of part I, we obtain (going back to G_R).

$$(2.5) \quad D_2('\nabla^p u, G_r) \leqq 2(1 - e_0/K)^{-1}(2p)! \cdot K^p \cdot L^2 \cdot (R-r)^{-2p}, \ \ L = |\, e\,|_{RK}{}^0,$$
$$u = u_0 \ \text{ or } \ u^*, \ \ 0 \leqq r < R, \ \ p = 0, 1, 2, \cdots, e_0 = 2.718 \cdots.$$

(2.5) yields bounds for all the derivatives of u involving at most one differentiation with respect to x^n. But on spheres interior to G_R, we may also apply Lemma 3.3 of part I *once* with $\gamma = n$ to each $'\nabla^p u$ and conclude that all the $'\nabla^p u_{x^n} \in \mathfrak{P}_2$ on such spheres. But then we obtain the equations

$$(2.6) \qquad '\nabla^p u_{x^n x^n} = -\sum_{\alpha=1}^{n-1} ('\nabla^p u_{x^\alpha x^\alpha} + '\nabla^p e_{\alpha x^\alpha}) - '\nabla^p e_{n x^n}$$

which shows that all the $'\nabla^p u_{x^n} \in \mathfrak{P}_2$ on each G_r with $r < R$, and, together with (2.5), yields the desired bounds on the $d^{**}{}_p(u, G_r)$.

The proofs of the second results for Q_{R0} and Q^*_R are practically identical with the proof of (ii) of Theorem 3.6, part I, and proceed as follows: We note that e satisfies

$$(2.7) \quad \begin{gathered} d_0[e - e(y_0), G(y_0, r)] \leqq L(r/\delta_{y_0})^{p+\mu}, \ \ |\, e(y_0)| \leqq Z_2(n,\mu) \cdot L \cdot \delta_{y_0}{}^{-p}, \\ L = Z_3(n,\mu) \cdot |\, e\,|_\mu{}^0, \ \ y_0 \in G_R. \end{gathered}$$

We shall show in both cases that (2.7) implies that

$$(2.8) \qquad d_1[u - l_{y_0 r}, G(y_0, r)] \leqq Z_4(n,\mu) \cdot L \cdot (r/\delta_{y_0})^{p+\mu}, \qquad 0 \leqq r \leqq \delta_y;$$

the result will then follow from Lemma 2.2 since y_0 is arbitrary.

We first consider Q^*_R and let y_0 be fixed. For each s, $0 \leqq s \leqq 1$, define $\psi(s) = L^{-1}$ times the sup of the left side of (2.8) $(u = u^*)$ for all $G(x_0, R)$ containing y_0, all e in $\mathfrak{L}_2$ and satisfying (2.7) (for all r) on $G(x_0, R)$, and $r = s\delta_{y_0}$. Then choose an arbitrary $G(x_0, R)$ containing y_0, an arbitrary

e in $\mathfrak{L}_2$ and satisfying (2.7) there (for all r), choose $0 < r \leqq r_0 \leqq \delta_{y_0}$, and let $u^* = Q^*_R(e)$. We write

$$(2.9) \qquad u^* = u^*_{y_0 r_0} + h^*_{y_0 r_0}, \qquad u^*_{y_0 r_0} = v^*_{y_0 r_0} - e_n(y_0) k^*_{r_0 y_0}$$

where $u^*_{y_0 r_0}$, $v^*_{r_0 y_0}$, and $k^*_{r_0 y_0}$ are the $G(y_0, r_0)$-*-quasi-potentials of e, $e - e(y_0)$, and e_0 where $e_{0\alpha} = -\delta_{\alpha n}$, respectively, and $h^*_{y_0 r_0}$ is the harmonic function $= u$ on $S^-(y_0, r_0)$. From the symmetry lemma, we see (by comparing U^* and $U^*_{y_0 r_0} = 0$ on $\partial\Gamma(y_0, r_0)$, etc.) that

$$(2.10) \qquad D_2[u^*, G(y_0, r_0)] = D_2[u^*_{y_0 r_0}, G(y_0, r_0)] + D_2[h^*_{y_0 r_0}, G(y_0, r_0)].$$

From (2.7), we conclude that

$$(2.11) \qquad d_0[e, G(y_0, r)] \leqq Z_5(n, \mu) \cdot (r/\delta_{y_0})^{\rho}.$$

Thus, using the symmetry, (2.10), (2.11), the fact that $B(y_0, r_0) \subset \Gamma(y_0, r_0)$, and the first result for Q^*_R, we see that

$$(2.12) \qquad d_1[h^*_{y_0 r_0}, B(y_0, r_0)] \leqq Z_6(n, \mu, \epsilon) \cdot (r_0/\delta_y)^{\rho - \epsilon}, \qquad 0 < \epsilon \leqq \rho.$$

Finally, it is easy to see that

$$(2.13) \quad k^*_{y_0 r_0}(x) = x^n + k^{**}_{y_0 r_0}(x),\ k^{**}_{y_0 r_0 x^n}(0, x'_n) = 0,\ \nabla k^{**}_{y_0 r_0}(0, y'_{0n}) = 0,$$

$k^{**}_{y_0 r_0}$ being the harmonic function on $\Gamma(y_0, r_0)$ which $= |x^n|$ on $\partial\Gamma(y_0, r_0)$.

Now, clearly,

$$(2.14) \quad \begin{aligned} d_1[u^* - l_{y_0 r}, G(y_0, r)] &\leqq d_1[v^*_{y_0 r_0} - l_r^1, G(y_0, r)] \\ &+ d_1[h^*_{y_0 r_0} - l_r^2, G(y_0, r)] + d_1[e_n(y_0) \cdot k^*_{y_0 r_0} - l_r^3, G(y_0, r)] \end{aligned}$$

where the various l's have their usual significance. From our definition of ψ, we have

$$(2.15) \qquad d_1[v^*_{y_0 r_0} - l_r^1, G(y_0, r)] \leqq L \cdot (r_0/\delta_y)^{\rho+\mu} \cdot \psi(r/r_0)$$

since $e - e(y_0)$ satisfies (2.7) on $G(y_0 r_0)$ with L replaced by $L \cdot (r_0/\delta_y)^{\rho+\mu}$. Using (2.12) and Lemma 3.4 of part I, we obtain

$$(2.16) \quad \begin{aligned} d_1[h^*_{y_0 r_0} - l_r^2, G(y_0, r)] &\leqq d_1[h^*_{y_0 r_0} - l_{y_0}{}^4, G(y_0, r)] \\ &\leqq d_1[h^*_{y_0 r_0} - l_{y_0}{}^4, B(y_0, r)] \\ \leqq (r/r_0)^{\rho+1} d_1[h^*_{y_0 r_0} - l_{y_0}{}^4, B(y_0, r_0)] &\leqq (r/r_0)^{\rho+1} d_1[h^*_{y_0 r_0}, B(y_0, r_0)] \\ \leqq Z_6(n, \mu, \epsilon)\, (r_0/\delta_y)^{\rho-\epsilon} (r/r_0)^{\rho-1} & \end{aligned}$$

where $l_{y_0}{}^4$ is tangent to $h^*{}_{y_0r_0}$ at y_0. Finally, from (2.13), we obtain

$$
\begin{aligned}
& d_1[e_n(y_0)\cdot k^*{}_{y_0r_0}-l_r{}^3, G(y_0,r)] \leqq |e_n(y_0)|\cdot d_1[k^{**}{}_{y_0r_0}-l_{y_0}{}^5, G(y_0,r)] \\
(2.17) \qquad & \leqq (r/r_0)^{\rho+1}\cdot|e_n(y_0)|\cdot d_1[k^{**}{}_{y_0r_0}, \Gamma(y_0,r_0)] \\
& \leqq (2a_n)^{\frac{1}{2}}\cdot Z_2\cdot L\cdot(r_0/\delta_{y_0})^{\rho}\cdot(r/r_0)^{\rho+1}, \qquad a_n = m[B(0,1)].
\end{aligned}
$$

Combining (2.14) to (2.17), setting $s=r/\delta_{y_0}$, $t=r_0/\delta_{y_0}$, combining the last two terms, and using the arbitrariness of $G(x_0,R)$ and e, we obtain

$$\psi(s) \leqq t^{\rho+\mu}\psi(s/t) + Z_7(n,\mu,\epsilon)t^{-\epsilon}(s/t)^{\rho+1}.$$

The result for $Q^*{}_R$ now follows from the last part of the proof of (ii), Theorem 3.6, part I.

The proof for Q_R is the same except that we may take $u_{0r_0y_0}=v_{0r_0y_0}$ in (2.9) and (2.10) follows since $u_{0r_0y_0}=0$ on $\partial G(y_0,r_0)$.

3. Regularity properties along $x^n=0$. In this section, we prove certain differentiability properties of the solutions of differential equations in integral form of the type

$$(3.1) \qquad \int_{G_R}\{w_{,\alpha}{}^i(a_{ij}{}^{\alpha\beta}u_{,\beta}{}^j+b_{ij}{}^{\alpha}u^j+e_i{}^{\alpha})+w^i(b^*{}_{ij}{}^{\alpha}u_{,\alpha}{}^j+c_{ij}u^j+f_i)\}dx=0, \qquad i=1,\cdots,N,$$

for all w in $\mathfrak{P}^*{}_{20}$ where $\mathfrak{P}^*{}_{20}$ consists of all vectors w in $\mathfrak{P}_2$ with

$$(3.2) \qquad w^i=0 \text{ on } S_R^-,\ i=1,\cdots,N, \qquad w^i=0 \text{ on } \sigma_R,\ i=1,\cdots,k, \qquad 0\leqq k\leqq N.$$

In general, we shall also assume that

$$(3.3) \qquad u^i=0 \text{ on } \sigma_R, \quad i=1,\cdots,k.$$

We assume that

$$(3.4) \qquad a(0)=a_0, \qquad a_{0ij}{}^{\alpha\beta}=\delta^{\alpha\beta}\cdot\delta_{ij}$$

and also that if x_0 is any point in G_R, there is a linear transformation of E^n into itself so that the new $'a(x_0)=a_0$. We assume at least that the a's are Lipschitz with b, b^*, and c bounded and measurable and e and $f\,\varepsilon\,\mathfrak{L}_2$. Additional assumptions will be made about the coefficients as desired. We do not assume that b and b^* are related in any way. The integer k above will be held fixed throughout this section.

As in part I, we assume that $u\,\varepsilon\,\mathfrak{P}$, and we wish to conclude further

differentiability properties of u by making further assumptions about the coefficients. As before, we write

$$u = u_0 + H \tag{3.5}$$

where now

$$u_0 \in \mathfrak{P}^*_{20},\ H^i = 0,\ i = 1, \cdots, k,\ H_{x^n}{}^i = 0,\ i = k+1, \cdots, n \text{ on } \sigma_R. \tag{3.6}$$

The harmonic vector H may be found by first extending u to B_R by

$$\begin{aligned} u^i(-x^n, x'_n) &= -u^i(x^n, x'_n), \quad i = 1, \cdots, k, \\ u^i(-x^n, x'_n) &= u^i(x^n, x'_n), \quad i = k+1, \cdots, n \end{aligned} \tag{3.7}$$

and then choosing H to coincide with u on ∂B_R; it is easy to see from the uniqueness that the H^i also satisfy (3.7) and hence (3.6). Then we note that

$$\int_{G_R} w_{x^\alpha}{}^i a_{0ij}{}^{\alpha\beta} H_{x^\beta}{}^j dx = 0 \text{ for all } w \in \mathfrak{P}^*_{20}. \tag{3.8}$$

Then reasoning as in Section 4 of part I, we see that

$$\begin{aligned} &u_0 = Tu_0 + V + W, \quad Tu_0 = Q_R[e_0(u_0)] + P_R[f_0(u_0)], \\ &V = Q_R[e_0(H)] + P_R[f_0(H)], \quad W = Q_R(e) + P_R(f) \\ &e_0(\phi) = (a - a_0) \cdot \nabla\phi + b \cdot \phi, \quad f_0(\phi) = b^* \cdot \nabla\phi + c \cdot \phi \end{aligned} \tag{3.9}$$

where $Q_R(e)$ and $P_R(e)$ are now the vectors $[Q_R{}^i(e)]$ and $[P_R{}^i(e)]$ where

$$\begin{aligned} &Q_R{}^i(e) = Q_{0R}(e_i{}^\alpha) \text{ and } P_R{}^i(f) = P_{0R}(f_i), \quad i = 1, \cdots, k, \\ &Q_R{}^i(e) = Q^*_R(e_i{}^\alpha) \text{ and } P_R{}^i(f) = P^*_R(f_i), \quad i = k+1, \cdots, N. \end{aligned} \tag{3.10}$$

The proof of Theorem 3.1 below is just like that of Theorem 4.1 of part I; the spaces and norms mentioned are defined in the obvious way:

THEOREM 3.1. (i) *If a is Lipschitz and b, b^*, and c are bounded and measurable with, say,*

$$\begin{aligned} &| a(x_1) - a(x_2) | \leqq L_1 \cdot | x_1 - x_2 |, \\ &\qquad | b(x) | \leqq L_2, \quad | b^*(x) | \leqq L^*_2, \quad | c(x) | \leqq L_3, \end{aligned} \tag{3.11}$$

then T is an operator on $\mathfrak{P}^*_{20}$ and on $\mathfrak{P}^*_{20\lambda}$, $\lambda = \rho - 1 + \mu$, $0 < \mu < 1$, where

$\| T \|^*_0 \leqq C_{10}(L_1, L_2, L^*_2, L_3, R_0) R,$

$$\| T \|^*_{0\lambda} \leqq C_{11}(n, \mu, L_1, L_2, L^*_2, L_3, R_0) \cdot R, R \leqq R_0$$

$$C_{10} = L_1 + C_1(L_2 + L^*_2) + C_1{}^2 R_0 L_3.$$

(ii) *If, also* $b \varepsilon C_\mu{}^0$ *with*

$$|b(x_1) - b(x_2)| \leqq L_4 \cdot |x_1 - x_2|^\mu, \quad 0 < \mu < 1, \tag{3.12}$$

then T *is an operator on* $C^*{}_{0\mu}{}^1$ *and*

$$\| T \|^*{}_{0\mu}{}^1 \leqq C_{12}(n; \mu, L_1, \cdots, L_4, L^*{}_2, R_0) \cdot R, R \leqq R_0.$$

(iii) *If, also, the coefficients are all analytic on* G_{R_0} *with*

$$|\nabla^p a| \leqq A \cdot p!F^p, \quad |\nabla^p b| \leqq B \cdot p!F^p, \quad |\nabla^p b^*| \leqq B^* p!F^p, \quad |\nabla^p c| \leqq Cp!F^p,$$

on G_{R_0} *for* $p = 0, 1, 2, \cdots$, *then* T *is an operator on* $\mathfrak{B}^*{}_{0RK}{}^1$ *with*

$$\| T \|^*{}_{0RK}{}^1 \leqq C_{13}(n, K, A, B, B^*, C, F, R_0) \cdot R \text{ if } R \leqq R_0, K > e = 2.718 \cdots.$$

COROLLARY. *If* R *is so small that* $\| T \|^*{}_0 < 1$, *then there exists a unique solution* u *of* (3.1), *in* $\mathfrak{P}_2$ *on* G_R *which coincides on* $S_R{}^-$ *with any given function* u^* *in* $\mathfrak{P}_2$ *on* G_R *and satisfies* (3.3).

For we may write (3.5) where H is determined by (3.6) and then solve (3.9) for u_0 in $\mathfrak{P}^*{}_{20}$.

THEOREM 3.2. *Suppose* a, b, b^*, *and* c *satisfy the hypotheses of Theorem* 3.1(i), *suppose* R *is small enough so that*

$$R(L_1 + C_1 L^*{}_2) = A < 1, \tag{3.13}$$

suppose u, e, *and* $f \varepsilon \mathfrak{Q}_2$ *on* G_R, *and suppose* $u \varepsilon \mathfrak{P}_2$ *and satisfies* (3.1) *and* (3.3) *on each* G_r *with* $r < R$. *Then*

$$d_1(u, G_r) \leqq d_1(U_0, G_R) + (2e_0)^{\frac{1}{2}} K d_0(U, G_R) \cdot (R - r)^{-1}, \quad (e_0 = 2.718 \cdots) \tag{3.14}$$

where

$$\begin{aligned} d_1(U_0, G_R) &\leqq (1 - A)^{-1}[d_0(e, G_R) + GRd_0(f, G_R) \\ &\qquad + (L_2 + C_1 R L_3) d_0(u, G_R)] \\ d_0(U_0, G_R) &\leqq C_1 R d_1(U_0, G_R), \quad d_0(U, G_R) \leqq d_0(u, G_R) + d_0(U_0, G_R), \\ K &\leqq 1 + AR \cdot (1 - A)^{-1}. \end{aligned} \tag{3.15}$$

Proof. The proof is identical with that of Theorem 4.2 of part I with G_r replacing B_r, $\mathfrak{P}^*{}_{20}$ replacing $\mathfrak{P}_{20}$, $S_r{}^-$ replacing S_r, etc. down to the step (4.16) where

$$D_2(U, G_r) \leqq K D_2(H_r, G_r). \tag{3.16}$$

Now, if we extend U and H_r to B_r and B_R by (3.7), we see that (3.16) holds with G_r replaced by B_r. The result follows from Theorem 3.4 of part I.

The proof of the following theorem is identical with that of Theorem 4.3 of part I except for the obvious changes (B_R replaced by G_R, numbering of equations, $\mathfrak{P}_2$ replaced by $\mathfrak{P}^*_{20}$, etc.):

THEOREM 3.3. *Suppose a, b, b^*, and c and a_p, b_p, b^*_p, and c_p satisfy the hypotheses of Theorem* 3.1(i) *uniformly in p, suppose R is small enough so that $C_{10}R < 1$, suppose $a_p(0) = a(0)$, suppose $a_p \to a$, etc., almost everywhere, and suppose $e_p \to e$ and $f_p \to f$ strongly in $\mathfrak{L}_2$ on G_R.*

(i) *If $u \varepsilon \mathfrak{P}_2$, satisfies* (3.3), *and is a solution of* (3.1) *on G_R and if u_p is, for each p, that solution in $\mathfrak{P}_2$ of* $(3.1)_p$ *and* (3.3) *which $= u$ on S_R^-, then $u_p \to u$ strongly in $\mathfrak{P}_2$ on G_R.*

(ii) *If, for each p, u_p is a solution in $\mathfrak{P}_2$ of* $(3.1)_p$ *and* (3.3) *on each G_r with $r < R$ and if $u_p \to u$ strongly in $\mathfrak{L}_2$ on G_R, then $u \varepsilon \mathfrak{P}_2$ and is a solution of* (3.1) *and* (3.3) *on each G_r with $r < R$ and $u_p \to u$ weakly in $\mathfrak{P}_2$ on each such G_r.*

We come now to our principal theorem on differentiability:

THEOREM 3.4. *Snppose that u is any solution of* (3.1) *and* (3.3) *which $\varepsilon \mathfrak{P}_2$ on G_R.*

(i) *Suppose that a, b, b^*, and c satisfy the hypotheses of Theorem* 3.1(i), *suppose that $C_{11}R < 1$ and $R \leqq R_0$, and suppose e and $f \varepsilon \mathfrak{L}_{2\lambda}$, $\lambda = \rho - 1 + \mu$, $0 < \mu < 1$, an B_R. Then $\nabla u \varepsilon \mathfrak{L}_{2\lambda}$ and $u \varepsilon C_\mu{}^0$ on B_R.*

(ii) *Suppose also that b and $e \varepsilon C_\mu{}^0$ with b satisfying* (3.12) *on G_R and suppose that $C_{12} \cdot R < 1$ and $R \leqq R_0$. Then $u \varepsilon C_\mu{}^1$ on G_R.*

(iii) *If a, b, and $e \varepsilon C_1{}^0$, a, b, b^*, and c satisfy the hypotheses of* (i), *and $f \varepsilon \mathfrak{L}_2$ on B_R, then $\nabla u \varepsilon \mathfrak{P}_2$ on each G_r with $r < R$.*

(iv) *If a, b, and $e \varepsilon C_\mu{}^{k-2}$ and b^*, c, and $f \varepsilon C_\mu{}^{k-3}$, $0 < \mu < 1$, $k \geqq 3$, then $u \varepsilon C_\mu{}^{k-1}$ on each G_r with $r < R$.*

(v) *If a, b, b^*, c, e, and f are analytic on G_{R_0}, then u is analytic on G_R for each sufficiently small R.*

Proof. In (i) and (ii) our assumptions guarantee that $\| T \| < 1$. From the symmetry properties (3.6) and (3.7) for H and from Theorems

3.4 and 3.5 of part I, it follows that V belongs to whatever space is being considered. This is also true of W, using the theorems of Section 2.

The proofs of (iii) and (iv) are like those of (iv) and (v) of Theorem 4.4 of part I, except that γ is kept $\leqq n-1$ when applying the device of Lichtenstein. We conclude from this that all the u_{x^γ} for $\gamma \leqq n-1$ belong to $\mathfrak{P}_2$ and satisfy (3.3) on each G_r with $r < R$. Moreover, on spheres interior to G_R, the device may be applied with $\gamma = n$ showing that $u_{x^n} \varepsilon \mathfrak{P}_2$ on such spheres. But then a simple Green's theorem for $\mathfrak{P}_2$ functions allows us to replace equations (3.1) by the corresponding system of differential equations (almost everywhere). These may be solved for the $u_{x^n x^n}{}^i$ obtaining

$$(3.17) \qquad u_{x^n x^n}{}^i = \sum_{\alpha=1}^{n-1} A_{nj}{}^{i\alpha} u_{x^n x^\alpha}{}^j + \sum_{\alpha,\beta=1}^{n-1} B_j{}^{i\alpha\beta} u_{x^\alpha x^\beta}{}^j + C_j{}^{i\alpha} u_{x^\alpha}{}^j + D_j{}^i u^j + F^i$$

where in (v) all the coefficients are analytic if R_0 is small enough, in (iii) the A's and B's $\varepsilon\, C_1{}^0$, C and D are bounded and measurable and $F \varepsilon\, \mathfrak{L}_2$, and in (iv) the A's and B's $\varepsilon\, C_\mu{}^{k-2}$ and C, D, and $F \varepsilon\, C_\mu{}^{k-3}$. The result (iii) follows immediately. To obtain (iv), we first apply the device of Lichtenstein with $\gamma \leqq n-1$ on the whole of each G_r as many times as the coefficients allow. Then by applying the device with $\gamma = n$ on the interior of G_r and using (3.17) and its derivatives, we obtain the desired results.

In the analytic case (v), we conclude from (iv) that $u \varepsilon\, C^\infty$ on each G_r with $r < R$. If we choose R so small that $R \leqq R_0$ and $C_{13} R < 1$, the argument for (i) and (ii) shows that $u \varepsilon\, \mathfrak{B}^*{}_{0RK}{}^1$ from which we conclude using Lemma 2.3 that all the $'\nabla^p u$ and $'\nabla^p u_{x^n}$ are continuous in x and analytic in x'_n on each such G_r. Our previous regularity results show that u is analytic interior to such G_r. The $\mathfrak{B}^*{}_{0RK}{}^1$ bounds, together with repeated differentiations of (3.17) suffice to obtain the necessary bounds for the derivatives of u; or a sort of dominating function method like that in the Cauchy-Kowalewsky theorem may be used.

4. Manifolds with boundary. For an n-dimensional Riemannian manifold $\mathfrak{M}$ with boundary $\mathfrak{B}$ of class $C_\mu{}^k$ $(0 \leqq \mu \leqq 1)$, (C^∞, analytic) we adopt the standard definition: each point of $\mathfrak{M}(U\mathfrak{B})$ is contained in some set $\mathfrak{N}$ open on $\mathfrak{M}$ which is either the homeomorphic image of the unit ball or of the part of it where $x^n \leqq 0$ in which latter case, the points where $x^n = 0$ correspond to $\mathfrak{N} \cap \mathfrak{B}$; any two overlapping coordinate systems are related by a transformation of class $C_\mu{}^k$ (C^∞, or analytic). An admissible coordinate system for such a manifold will be any homeomorphism of a Lipschitzian

domain in G in E^n onto a set $\mathfrak{N}$ open on $\mathfrak{M}$ which is related to the "preferred," coordinate systems above by a transformation of class $C_\mu{}^k$ (C^∞, or analytic); if $\mathfrak{N} \cap \mathfrak{B}$ is not empty, G must lie in the half-plane $x^n < 0$ and the part of ∂G on $x^n = 0$ must correspond to $\mathfrak{N} \cap \mathfrak{B}$. In any admissible coordinate system, the g_{ij} are of class $C_\mu{}^{k-1}$.

As in part I, we shall assume that $\mathfrak{M}$ is at least of class $C_1{}^1$, in which case the g_{ij} are merely Lipschitzian. We shall be concerned with exterior differential forms on $\mathfrak{M}$. Since we have not required (and shall not) $\mathfrak{M}$ to be orientable, we shall consider both *even* and *odd* forms (see [2], § 3) on $\mathfrak{M}$. The law of transformation of the components under coordinate transformations is

$$
\begin{aligned}
&{}'\omega_{i_1\cdots i_r}(x') = \epsilon \sum_{j_1<\cdots<j_r} \omega_{j_1\cdots j_r}[x('x)] \frac{\partial(x^{j_1},\cdots,x^{j_r})}{\partial('x^{i_1},\cdots,'x^{i_r})}, \\
&\epsilon = \begin{cases} +1, & \text{for even forms,} \\ J/|J|, & \text{for odd forms.} \end{cases} \qquad J = \frac{\partial(x^1,\cdots,x^n)}{\partial('x^1,\cdots,'x^n)}.
\end{aligned}
\tag{4.1}
$$

The differentiability class of a form ω is that of its components in all coordinate systems; on a manifold of class $C_\mu{}^k$, it is clear that no form can be of class $\geqq C_\mu{}^{k-1}$ (although in particular coordinate systems the components might have higher class). Forms of class $\mathfrak{L}_2$ are defined as usual and the $\mathfrak{L}_2$ inner product of two (both even or both odd) forms ω and η and the norm of ω will be denoted by (ω, η) and $|\omega|$, respectively. Forms of class $\mathfrak{P}_2$ were defined in part I for manifolds without boundary; it is clear that the definition there given carries over to the present case along with those of the $\mathfrak{P}_2$ inner product and norm $((\omega, \eta))$ and $\|\omega\|$, depending on a finite number $\mathfrak{N}$ of admissible coordinate systems covering $\mathfrak{M}$. The differential operators d and δ and the Dirichlet integral are defined as usual (see part I, § 5). The dual operator $*$ is defined as usual (see, for instance [3], p. 129 or [2]).

There are parallel theories for even and odd forms. From now on we assume that all forms are of some one kind.

We have immediately the theorem:

THEOREM 4.1. *The spaces $\mathfrak{L}_2$ and $\mathfrak{P}_2$ of even or odd forms are Hilbert spaces with the norms above. The operators d and δ are bounded operators from $\mathfrak{P}_{2e}$ to $\mathfrak{L}_{2e}$ and from $\mathfrak{P}_{20}$ to $\mathfrak{L}_{20}$ and the Dirichlet integral $D(\omega)$ is lower-semicontinuous with respect to weak convergence in either $\mathfrak{P}_{2e}$ or $\mathfrak{P}_{20}$. Finally if $\omega_p \to \omega$ weakly in $\mathfrak{P}_{2e}$ (or $\mathfrak{P}_{20}$) then $\omega_p \to \omega$ strongly in $\mathfrak{L}_{2e}$ (or $\mathfrak{L}_{20}$).*

We shall be concerned with the boundary values of forms. We begin with the following definition:

Definition 4.1. By the *boundary values* of a form on $\mathfrak{M}(U\mathfrak{B})$, we merely mean that form restricted to $\mathfrak{B}$; that is the boundary value $b\omega$ of ω is given by

$$(4.2) \qquad b\omega = \sum_{i_1<\cdots<i_r\leqq n} \omega_{i_1\cdots i_r}(0, x'_n)\, dx^{i_1}\cdots dx^{i_r}$$

whenever $\omega_{i_1\cdots i_r}(x)$ are the components of ω in an admissible boundary coördinate system. The boundary value of ω is said to be *of class* $\mathfrak{L}_2$ *or* $\mathfrak{P}_2$ *along* $\mathfrak{M}$, if $\mathfrak{M}$ is of class C_1^1, or of class $C_\nu^l \leqq C_\mu^{k-1}$ (C_μ^k for 0-forms) if $\mathfrak{M}$ is of class C_μ^k, or C^∞, or analytic, if and only if its components under any admissible boundary coordinate system have the indicated class as functions of x'_n. If ω is of class $\mathfrak{P}_2$ on $\mathfrak{M}$, it is understood that the components of ω are to be replaced by the corresponding functions of Theorem 2.7 of part I as in Theorems 2.9, etc., of part I.

Before proceeding further, we prove the following lemma:

LEMMA 4.1. (a) *If* $\omega \varepsilon \mathfrak{P}_2$ *on* $\mathfrak{M}$, *its boundary value* $\varepsilon \mathfrak{L}_2$ *on* $\mathfrak{B}$.

(b) *If the boundary value of some form* ω_0 *is of class* $\mathfrak{P}_2$ *along the boundary* $\mathfrak{B}$ *of a manifold* $\mathfrak{M}$ *of class* C_1^1, *then there is a form* $\omega \varepsilon \mathfrak{P}_2$ *on* $\mathfrak{M}$ *with the same boundary value almost everywhere on* $\mathfrak{M}$. *If, also* $d\omega_0(\delta\omega_0) \varepsilon \mathfrak{P}_2$ *along* $\mathfrak{B}$, *we may choose* ω *so that* $d\omega(\delta\omega) \varepsilon \mathfrak{P}_2$ *on* $\mathfrak{M}$.

(c) *If* $\mathfrak{M}$ *is of class* C_μ^k (C^∞) *and the boundary value of* ω_0 *is of class* C_μ^{k-1} (C_μ^k *for* 0 *forms*) (C^∞) *along* $\mathfrak{B}$, *then there is a form* ω *of class* C_μ^{k-1} (C_μ^k *for* 0 *forms*) (C^∞) *on* $\mathfrak{M}$ *with the same boundary value. If we also assume* $d\omega(\delta\omega)$ *of class* C_μ^{k-1} (*on* $\mathfrak{M}$ *of class* C_μ^k) *we may choose* ω *so that* $d\omega(\delta\omega) \varepsilon C_\mu^{k-1}$ *on* $\mathfrak{M}$.

(d) *If* $\mathfrak{M}$ *is analytic and the boundary value of* ω_0 *is analytic along* $\mathfrak{B}$, *then* ω *may be chosen to be of class* C^∞ *on* $\mathfrak{M}$ *and analytic in some neighborhood of any given boundary point.*

Proofs. (a) follows from Theorem 2.9 of part I. We prove (b) and (c) simultaneously as follows: We may cover B with the ranges $\mathfrak{R}_1, \cdots, \mathfrak{R}_Q$ of admissible boundary coordinate systems whose domains are all the part $x^n \leqq 0$ of the unit ball $B(0;1)$, and we may find a partition of unity $\phi_1, \cdots, \phi_S$, each of whose functions is of class $C_1^1(C_\mu^k)$, and such that each function whose support intersects $\mathfrak{B}$ has support in some one $\mathfrak{R}_q$. Since the

sum of all those $\phi_s(P)$ whose support intersects $\mathfrak{B}$ is 1 on $\mathfrak{B}$, we may define $\omega = \omega_1 + \cdots + \omega_S$ where each $\omega_s = 0$ unless the support of ϕ_s intersects $\mathfrak{B}$ in which case we define $\omega_s = 0$ outside $\mathfrak{N}_q$ and define ω_s in $\mathfrak{N}_q$ by

$$\omega^{(q)}{}_{si_1\cdots i_r}(x^n, x'_n) = \omega^{(q)}{}_{0i_1\cdots i_r}(0, x'_n) \cdot \phi_s{}^{(q)}(x^n, x'_n). \tag{4.2}$$

The differentiability results, including those about $d\omega$ follows from (4.2); to get the ones about $\delta\omega$, begin by taking duals. In the analytic case we make sure that one of the $\mathfrak{N}_q$ is a boundary coordinate system about the given point P and that one $\phi_s \equiv 1$ in some neighborhood of P.

Definition 4.2. Suppose ω is defined on $\mathfrak{M}$. Then we define the *tangential* and *normal parts* $t\omega$ and $n\omega$ of its boundary value by the condition that

$$t\omega = \sum_{i_1 < \cdots < i_r < n} \omega_{i_1\cdots i_r} dx^{i_1} \cdots dx^{i_r}, \quad n\omega = b\omega - t\omega, \quad \text{on } x^n = 0 \tag{4.3}$$

in any admissible boundary coordinate system in which

$$\omega = \sum_{i_1 < \cdots < i_r \leqq n} \omega_{i_1\cdots i_r} dx^{i_1} \cdots dx^{i_r} \text{ on } x^n = 0. \tag{4.4}$$

Since $\partial x^n / \partial' x^i = 0$ on $'x^n = 0$ in the relation between two overlapping admissible boundary coordinate systems, it follows that $t\omega$ and hence $n\omega$ is invariantly defined on $\mathfrak{B}$.

The following lemma is important:

LEMMA 4.2. (a) *If $\omega_p \to \omega$ weakly in $\mathfrak{P}_2$ on $\mathfrak{M}$, then $b\omega_p \to b\omega$, $t\omega_p \to t\omega$, and $n\omega_p \to n\omega$ strongly in $\mathfrak{L}_2$ on $\mathfrak{B}$.*

(b) *If $\omega \varepsilon \mathfrak{P}_2$ on $\mathfrak{M}$, then $t(*\omega) = *(n\omega)$ and $n(*\omega) = *(t\omega)$ and $*(b\omega) = b(*\omega)$ on $\mathfrak{B}$. Here $*(n\omega)$ and $*(t\omega)$ may be found by first extending $n\omega$ and $t\omega$ to $\mathfrak{M}$ in any way, taking the duals, and then finding the boundary values; the result is independent of the way $n\omega$ and $t\omega$ are extended.*

Proofs. (a) follows from Theorems 2.11 and 2.12 of part I. From our strong convergence theorems, it suffices to prove (b) for Lipschitz forms. Let ω be such a form and let P be any point of $\mathfrak{B}$. There exists an admissible boundary coordinate system with domain B_R which carries the origin into P with $g_{ij}(0) = \delta_{ij}$. In that coordinate system, we have (see formula 1.10) of [3], for instance)

$$(*\omega)_{i_1\cdots i_{n-r}} = e_{j_1\cdots j_r i_1\cdots i_{n-r}} \omega_{j_1\cdots j_r} \quad (j \text{ not summed}) \text{ at } x = 0; \tag{4.5}$$

the j's being those positive integers $\leqq n$ which are not among the i's. The results follow from (4.5) and the definitions.

If $\mathfrak{M}$ is of class C''' and ϕ and ψ are forms of class C'' of the same kind and of degrees p and $(p+1)$, respectively, the following formula is derived in [3], § 2, from Stokes formula:

$$(d\phi, \psi) - (\phi, \delta\psi) = \int_{\mathfrak{B}} \phi \wedge * \psi. \tag{4.6}$$

From this, it follows that if $\mathfrak{M}$ is of class C'''', ω is of class C''' and ζ is of class C'', then

$$(d\omega, d\zeta) + (\delta\omega, \delta\zeta) = (\Delta\omega, \zeta) + \int_{\mathfrak{B}} (\zeta \wedge * d\omega - \delta\omega \wedge * \zeta), \qquad \Delta = \delta d + d\delta, \tag{4.7}$$

Δ being the Laplacian operator as used in [2] and [3] (see [3], p. 130). In Euclidean space with Cartesian coordinates, this is the negative of the ordinary Laplacian. In this connection, we note that if $\mathfrak{M} = G_R$, the part of the sphere $|x| < R$ for which $x^n < 0$, then

$$\begin{aligned} D_0(\omega) = 2 \int_{\sigma_R} (-1)^{n-r} \sum_{i_1 < \cdots i_r < n} \sum_{k=1}^{r} (-1)^k [\omega_{i_1 \cdots i_r} \omega_{i_1 \cdots i'_k \cdots i_r n x^{i_k}} \\ - \omega_{i_1 \cdots i_r x^{i_k}} \omega_{i_1 \cdots i'_k \cdots i_r n}] dx_1 \cdots dx_{n-1} \\ + \int_{G_R} \sum_{i_1 < \cdots < i_r} \sum_{\alpha=1}^{n} (\omega_{i_1 \cdots i_r x^\alpha})^2 dx \end{aligned} \tag{4.8}$$

for any form ω of class C'' which is zero on and near the spherical surface of G_R, $D_0(\omega)$ being the Dirichlet integral $D(\omega)$ referred to Cartesian coordinates and σ_R being the part of G_R for which $x^n = 0$.

LEMMA 4.3. (a) *If ω (considered as a set of functions) $\varepsilon\, \mathfrak{P}_2$ on G_R, ω is zero on and near the spherical part of the surface of G_R, and if either $t\omega = 0$ or $n\omega = 0$, there exists a sequence ω_p of forms of class C^k for any desired k, on G_R which converge strongly in $\mathfrak{P}_2$ on G_R to ω and such that $t\omega_p = 0$ or $n\omega_p = 0$ (respectively) for each p.*

(b) *If ω satisfies the hypotheses of* (a), *then*

$$D_0(\omega) = \int_{G_R} \sum_{(i)\alpha} \omega_{(i) x^\alpha})^2\, dx. \tag{4.9}$$

Proof. Clearly (b) follows from (a) and (4.8). Also, since the condition $t\omega = 0$ is just the same as saying that the ω's with $i_r < n$ vanish on σ_R and $n\omega = 0$ is the same as saying that those with $i_r = n$ vanish on σ_R, part (a) is just reduced to proving the theorem for functions. If the

function ω is not required to be zero on σ_R, we extend ω to the whole of B_R by $\omega(-x^n, x'_n) = \omega(x^n, x'_n)$ and then note that the first spherical h-averages of ω are of class C' and have support $\subset B_R$ if h is small enough; these tend strongly in $\mathfrak{P}_2$ to ω on B_R. We may then repeat the process $k-1$ more times. If ω is zero on σ_R, we begin by extending ω to B_R by $\omega(-x^n, x'_n) = -\omega(x^n, x'_n)$ and then proceeding as above; we note that each of the successive averages vanishes on σ_R.

This lemma and equation (4.6) for smooth forms and manifolds allows us to prove the following important facts:

LEMMA 4.4. *Suppose α and β are any forms $\varepsilon\, \mathfrak{P}_2$ which are of the same kind and of proper degrees. Then $(\delta\alpha, d\beta) = 0$ and $(\alpha, d\beta) = (\delta\alpha, \beta)$, if either $n\alpha$ or $t\beta = 0$.*

Proof. We may select a finite covering of $\mathfrak{M}$ by coordinate neighborhoods $\mathfrak{N}_1, \cdots, \mathfrak{N}_Q$ covering $\mathfrak{M}$, the coordinate system corresponding to any boundary neighborhood being an admissible boundary coordinate system. It is clear that there is a number $R_0 > 0$ such that any geodesic sphere on $\mathfrak{M}$ of radius R_0 is contained in some one $\mathfrak{N}_q$. A finite number of new neighborhoods, each part of a sphere of radius $R_0/3$, also covers $\mathfrak{M}$. If we choose a partition of unity $\phi_1 + \cdots + \phi_S \equiv 1$, each function of which is Lipschitz and has support in one of these small neighborhoods we see that

$$(\delta\alpha, d\beta) = \sum_{s,t} (\delta\alpha_s, d\beta_t); \qquad \alpha_s = \phi_s\alpha, \beta_t = \phi_t\beta$$

where the sum is extended over all ordered pairs (s, t) such that the supports of α_s and β_t intersect. But for any such pair, the union of the supports is included in some one $\mathfrak{N}_q$ so our problem is reduced to that case. But for inner neighborhoods the theorem has been proved in part I, Lemma 7.1. But the same proof extends to boundary neighborhoods using Lemma 4.3.

LEMMA 4.5. (Gaffney [5]). *With each point P of $\mathfrak{M}$ and each $\epsilon > 0$ is associated an admissible coordinate system $\mathfrak{C}$ with domain G and range $\mathfrak{U}$ and a constant l such that*

$$D(\omega) \geqq (1-\epsilon) \int_G \sum_{i,\alpha} (\omega_{ix^\alpha})^2\, dx - l(\omega, \omega)$$

for any form $\omega\, \varepsilon\, \mathfrak{P}_2$ with support on $\mathfrak{U}$ and either $t\omega = 0$ or $n\omega = 0$ on $\mathfrak{B}$.

Proof. This has been proved for interior points in part I, Lemma 5.1. If P is a boundary point, it is clear that we may choose an admissible

boundary coordinate system with domain $G_R \cup \sigma_R$ which carries the origin into P and $G_R \cup \sigma_R$ into a neighborhood of P in which $g_{ij}(0) = \delta_{ij}$. Then, exactly as in the proof for interior points, we conclude that we may choose R so small that

$$D(\omega) \geqq D_0(\omega) - \epsilon \int_{G_R} \sum_{i,\alpha} (\omega_{ix^\alpha})^2 \, dx - l(\omega, \omega)$$

for some l and all $\omega \varepsilon \mathfrak{P}_2$ with support in $\mathfrak{N}$, $D_0(\omega)$ having its significance in Lemma 4.3. The result follows from that lemma.

The following theorem follows from the lemma above in exactly the same way as in the case of part I, Theorem 5.4.

THEOREM 4.2. *For each finite system $\mathfrak{N}$ of admissible coordinate systems whose ranges cover $\mathfrak{M}$, there are constants k and l such that*

$$D(\omega) \geqq k \| \omega \|^2 - l(\omega, \omega) \qquad (k > 0)$$

for any form in $\mathfrak{P}_2$ with either $t\omega = 0$ or $n\omega = 0$ on $\mathfrak{B}$, the norm being that corresponding to $\mathfrak{N}$.

5. Potentials; the decomposition theorem. In this and the next section, we shall assume that all of our forms are of the same kind, completely parallel theories being obtained for each kind.

Definition 5.1. We define the *closed linear manifolds $\mathfrak{P}_2^+$ and $\mathfrak{P}_2^-$* (see Theorem 4.2) of $\mathfrak{P}_2$ as the totality of forms in $\mathfrak{P}_2$ for which $n\omega = 0$ and $t\omega = 0$ on $\mathfrak{B}$, respectively.

Just as in part I, Section 6, we obtain the following result:

LEMMA 5.1. *Let M be any closed linear manifold of $\mathfrak{L}_2$ such that $M \cap \mathfrak{P}_2^+(\mathfrak{P}_2^-)$ is not empty. Then there exists a form ω in $M \cap \mathfrak{P}_2^+(\mathfrak{P}_2^-)$ which minimizes $D(\omega)$ among all such forms with $(\omega, \omega) = 1$.*

THEOREM 5.1. *The manifold $\mathfrak{H}^+(\mathfrak{H}^-)$ of harmonic fields in $\mathfrak{P}_2^+(\mathfrak{P}_2^-)$ is finite dimensional.*

THEOREM 5.2. *If $\omega \varepsilon \mathfrak{P}_2^+(\mathfrak{P}_2^-)$ and is $\mathfrak{L}_2$-orthogonal to $\mathfrak{H}^+(\mathfrak{H}^-)$, then there are positive constants λ^+ and λ^- for each $\mathfrak{N}$ such that*

$$D(\omega) \geqq \lambda^+ \| \omega \|^2 (\lambda^- \| \omega \|^2).$$

THEOREM 5.3. *If $\eta \varepsilon \mathfrak{L}_2$ and is $\mathfrak{L}_2$-orthogonal to $\mathfrak{H}^+(\mathfrak{H}^-)$, there is a unique form $\Omega^+(\Omega^-)$ in $\mathfrak{P}_2^+ \perp \mathfrak{H}^+ (\mathfrak{P}_2^- \perp \mathfrak{H}^-)$ such that*

$$(d\Omega^+, d\zeta) + (\delta\Omega^+, \delta\zeta) = (\eta, \zeta), \zeta \varepsilon \mathfrak{P}_2^+(\mathfrak{P}_2^-). \tag{5.1}$$

Definition 5.2. The functions Ω^+ and Ω^- are called the plus-potential and minus-potential of η, respectively.

The defining equations (5.1) for the potentials are a special case of the more general equations

$$(5.2) \qquad (d\omega - \phi, d\zeta) + (\delta\omega - \psi, \delta\zeta) - (\eta, \zeta) = 0, \ \zeta \in \mathfrak{P}_2^+ \text{ or } \mathfrak{P}_2^-$$

which were discussed in Section 7, part I. The differentiability results for such equations on the interior of $\mathfrak{M}$ follow from the discussion there given. But now, suppose we select a point P on $\mathfrak{B}$ and choose an admissible boundary coordinate system with domain G_R and range a boundary neighborhood $\mathfrak{U}$ of P such that $g_{ij}(0) = \delta_{ij}$. In such a system the conditions $n\omega = n\zeta = 0$ for $\mathfrak{P}_2^+$ and $t\omega = t\zeta = 0$ for $\mathfrak{P}_2^-$ correspond under a proper ordering of the sets (i) and (j) to the equations (3.3). Accordingly as in Section 7, part I, we see that if the support of ζ is confined to $\mathfrak{U}$, the system (5.2) reduces to the system (3.1) and (3.3) discussed in Section 3. Since the theorems of Section 3 parallel exactly those of part I, Section 4, we may conclude that the differentiability results for the plus and minus potentials stated near the beginning of Section 7, part I, hold right up to the boundary. We now extend these results as in Section 7, part I, and summarize as follows:

THEOREM 5.4. *Suppose* $\omega \in \mathfrak{L}_2 \ominus \mathfrak{H}^+ (\mathfrak{L}_2 \ominus \mathfrak{H}^-)$ *and* Ω *is its plus (minus)-potential.*

(i) *If* $\mathfrak{M}$ *is of class* C_1^1, *then* Ω, $d\Omega$, *and* $\delta\Omega$ *are in* $\mathfrak{P}_2^+(\mathfrak{P}_2^-)$.

(ii) *If* $\mathfrak{M}$ *is of class* C_1^1 *and* ω *is in* $\mathfrak{L}_{2\lambda}$ *with* $\lambda = \rho - 1 + \mu \, (\rho = n/2)$, *then* Ω, $d\Omega$, *and* $\delta\Omega$ *are in* $\mathfrak{P}_{2\lambda}^+(\mathfrak{P}_{2\lambda}^-)$ *and* C_μ^0 *if* $0 < \mu < 1$.

(iii) *If* $\mathfrak{M}$ *is of class* C_μ^k *and* $\omega \in C_\mu^{k-2}$ $(k \geqq 2, 0 < \mu < 1)$, *then* Ω, $d\Omega$, *and* $\delta\Omega \in C_\mu^{k-1}$. *If* $k \geqq 3$ *and* $\omega \in C_\mu^{k-3}$, *then* $\Omega \in C_\mu^{k-1}$.

(iv) *If* $\mathfrak{M}$ *and* ω *are of class* C^∞ *or analytic, then so is* Ω.

(v) *If* Ω *and* ω *are 0-forms, then* Ω *has an additional degree of differentiability in all cases above except the second half of* (iii).

In all cases, if we set $\alpha = d\Omega$ *and* $\beta = \delta\Omega$,

$$(5.3) \qquad \delta\alpha + d\beta = \delta(d\Omega) + d(\delta\Omega) = \omega, \qquad d\alpha = \delta\beta = 0;$$

$$(5.4) \quad (d\alpha, d\zeta) + (\delta\alpha - \omega, \delta\zeta) = (d\beta - \omega, d\zeta) + (\delta\beta, \delta\zeta) = 0, \ \zeta \in \mathfrak{P}_2^+(\mathfrak{P}_2^-).$$

Proof. The results for Ω follows directly from the discussion above and

Section 3. The proof of the results for $d\Omega$ and $\delta\Omega$ is like that of Theorem 7.1 of part I where it is already done for the interior of $\mathfrak{M}$. We choose a boundary point P and an admissible boundary coordinate system of the type described in the preceding paragraph and approximate (if necessary) to ω and the g_{ij} by smooth functions. For each of the approximating functions Ω, we see from formula (4.7) that $\Delta\Omega = \omega$ and

$$(5.5) \qquad t^* d\Omega = 0 \text{ if } \Omega \varepsilon \mathfrak{P}_2^+ \text{ and } t\delta\Omega = 0 \text{ if } \Omega \varepsilon \mathfrak{P}_2^-$$

since the integral over $\mathfrak{B}$ depends only on the tangential parts of each factor in each term and $t\zeta$ is arbitrary if $\zeta \varepsilon \mathfrak{P}_2^+$ and $n\zeta$ or $t^*\zeta$ is arbitrary if $\zeta \varepsilon \mathfrak{P}_2^-$. From Stokes' theorem we see that $td\phi = 0$ whenever $t\phi = 0$ and ϕ and $\mathfrak{M}$ are differentiable. From Lemma 4.2(b) and the formula for δ, it follows from this that

$$(5.6) \qquad n\phi = 0 \to t^*\phi = 0 \to td^*\phi = 0 \to n^*d^*\phi = 0 \to n\delta\phi = 0.$$

Hence, from this and (5.5), we see that both α and $\beta \varepsilon \mathfrak{P}_2^+(\mathfrak{P}_2^-)$ if ω and $\Omega \varepsilon \mathfrak{P}_2^+(\mathfrak{P}_2^-)$ at each stage of the approximation, that α and β satisfy (5.3), and hence (5.4) using Lemma 4.4. The approximation may then be carried through as before on each G_r with $r < R$. Since a finite number of the smaller boundary neighborhoods cover $\mathfrak{B}$, the results (5.4) for all ζ in $\mathfrak{P}_2^+(\mathfrak{P}_2^-)$ follow and the differentiability of α and β now follow from Section 3.

Remark. Except in the case of zero forms Ω, the individual derivatives of the individual components of Ω *do not*, in general, have the same differentiability properties as do $d\Omega$ and $\delta\Omega$ (the coordinate transformations will not allow it).

The following two theorems are useful and important:

THEOREM 5.5. *Suppose $\eta \varepsilon \mathfrak{P}_2$, H^+ and H^- are its projections in $\mathfrak{H}^+$ and $\mathfrak{H}^-$ and Ω^+ and Ω^- are the plus and minus potentials of $\eta - H^+$ and $\eta - H^-$, respectively, and $\alpha^\pm = d\Omega^\pm$, $\beta^\pm = \delta\Omega^\pm$. Then*

(i) *α^+ is the plus potential of $d\eta$ and β^- is the minus potential of $\delta\eta$ and $d\eta \varepsilon \mathfrak{L}_2 \ominus \mathfrak{H}^+$ and $\delta\eta \varepsilon \mathfrak{L}_2 \ominus \mathfrak{H}^-$.*

(ii) *If $\eta \varepsilon \mathfrak{P}_2^+$, then β^+ is the plus potential $\delta\eta \varepsilon \mathfrak{L}_2 \ominus \mathfrak{H}^+$.*

(iii) *If $\eta \varepsilon \mathfrak{P}_2^-$, then α^- is the minus potential of $d\eta \varepsilon \mathfrak{L}_2 \ominus \mathfrak{H}^-$.*

Proof. These results follow from Theorem 5.4, equation (5.4), and Lemma 4.4.

THEOREM 5.6. (i) *If* $\eta^+ \varepsilon \mathfrak{P}_2^+$ *and* $\eta^- \varepsilon \mathfrak{P}_2^-$, *there are unique forms* α *and* β, *where* $\alpha \varepsilon \mathfrak{P}_2^+$ *and is* $\mathfrak{L}_2$*-orthogonal to* $\mathfrak{H}^+$ *and* $\beta \varepsilon \mathfrak{P}_2^-$ *and is* $\mathfrak{L}_2$*-orthogonal to* $\mathfrak{H}^-$ *such that* $\delta\alpha = \delta\eta^+$, $d\alpha = 0$, $d\beta = d\eta^-$, $\delta\beta = 0$.

(ii) *If* $\eta \varepsilon \mathfrak{P}_2$, *there are unique forms* $\gamma \varepsilon \mathfrak{P}_2^+ \cap (\mathfrak{L}_2 - \mathfrak{H}^+)$ *and* ϵ *in* $\mathfrak{P}_2^- \cap (\mathfrak{L}_2 - \mathfrak{H}^-)$ *such that* $d\gamma = d\eta$, $\delta\gamma = 0$, $\delta\epsilon = \delta\eta$, $d\epsilon = 0$.

Proof. The uniqueness is evident. To prove (i), let Ω^+ and Ω^- be the respective plus and minus potentials of $\eta^+ - H^+$ and $\eta^- - H^-$ and let $\Gamma = \delta\Omega^+$ and $E = d\Omega^-$. From Theorem 5.5, we see that Γ is the plus potential of $\delta\eta^+$ and E is the minus potential of $d\eta^-$. Then, from Theorem 5.4 we conclude that $\alpha = d\Gamma$ and $\beta = \delta E$ have the desired properties. To prove (ii), let Ω^+ and Ω^- be the respective plus and minus potentials of $\eta - H^+$ and $\eta - H^-$ and let

$$A = d\Omega^+, \quad B = \delta\Omega^-, \quad \gamma = \delta A, \quad \epsilon = dB$$

and (ii) follows from Theorem 5.5(i).

Definition 5.3. We define the *linear sets* $\mathfrak{C}$ *and* $\mathfrak{D}$ as the sets of all forms of the form $\delta\alpha$ and $d\beta$, where $\alpha \varepsilon \mathfrak{P}_2^+$ and $\beta \varepsilon \mathfrak{P}_2^-$, respectively.

We now can prove an analog for the Kodaira decomposition theorem [6] for manifolds with boundary.

THEOREM 5.7. *The sets* $\mathfrak{C}$ *and* $\mathfrak{D}$ *and the set* $\mathfrak{H}$ *of all harmonic fields in* $\mathfrak{L}_2$ *on* $\mathfrak{M}$ *are closed linear manifolds in* $\mathfrak{L}_2$ *and*

$$\mathfrak{L}_2 = \mathfrak{C} \oplus \mathfrak{D} \oplus \mathfrak{H}. \tag{5.7}$$

Moreover, if $\omega \varepsilon \mathfrak{P}_2$, *its* $\mathfrak{L}_2$ *projections* γ, ϵ, *and* H *on* $\mathfrak{C}$, $\mathfrak{D}$, *and* $\mathfrak{H}$ *belong to* $\mathfrak{P}_2^+$, $\mathfrak{P}_2^-$, *and* $\mathfrak{P}_2$, *respectively, and* $\delta\gamma = d\epsilon = 0$.

Proof. That $\mathfrak{C}$ and $\mathfrak{D}$ are closed linear manifolds follows immediately from Theorems 5.6 and 5.2 and that $\mathfrak{H}$ is also follows from the theorems of part I, Sections 4 and 7. Using Lemma 4.4, we see that $\mathfrak{C}$ and $\mathfrak{D}$ are orthogonal and that $\mathfrak{C}$ and $\mathfrak{D}$ are both orthogonal to $\mathfrak{H} \cap \mathfrak{P}_2$ and, in fact, if $H \varepsilon \mathfrak{P}_2 \cap (\mathfrak{L}_2 \ominus \mathfrak{C} \ominus \mathfrak{D})$, then $H \varepsilon \mathfrak{H}$ ($\mathfrak{P}_2^+$ and $\mathfrak{P}_2^-$ are both everywhere dense in $\mathfrak{L}_2$).

Now, suppose $\eta \varepsilon \mathfrak{P}_2$ and let γ and ϵ be its projections on $\mathfrak{C}$ and $\mathfrak{D}$, respectively. Using Theorem 5.6 we conclude the existence of unique forms α and β in $\mathfrak{P}_2^+ \cap (\mathfrak{L}_2 \ominus \mathfrak{H}^+)$ and $\mathfrak{P}_2^- \cap (\mathfrak{L}_2 \ominus \mathfrak{H}^-)$ respectively, such that

$$\delta\alpha = \gamma, \quad d\alpha = 0, \quad \delta\beta = 0, \quad d\beta = \epsilon. \tag{5.8}$$

Since γ and ϵ are the projections of η on $\mathfrak{C}$ and $\mathfrak{D}$, we see from (5.8) that α and β satisfy

$$(5.9)\quad \begin{aligned} &(d\alpha, d\zeta^+) + (\delta\alpha, \delta\zeta^+) - (\eta, \delta\zeta^+) = (d\alpha, d\zeta^+) + (\delta\alpha, \delta\zeta^+) - (d\eta, \zeta^+) = 0 \\ &(d\beta, d\zeta^-) + (\delta\beta, \delta\zeta^-) - (\eta, d\zeta^-) = (d\beta, d\zeta^-) + (\delta\beta, \delta\zeta^-) - (\delta\eta, \zeta^-) = 0 \end{aligned}$$

for all ζ^+ in $\mathfrak{P}_2^+$ and ζ_2^- in $\mathfrak{P}_2^-$. Thus α is the plus potential of $d\eta$ and β is the minus potential of $\delta\eta$. The results follow from Theorems 5.3 and 5.5.

The following theorem contains further information concerning the decomposition (5.7):

THEOREM 5.8. *Suppose $\omega \in \mathfrak{L}_2$ and γ, ϵ, and H are its projections on $\mathfrak{C}$, $\mathfrak{D}$, and $\mathfrak{H}$, respectively, and suppose Ω^+ and Ω^- are the plus and minus potentials of $\omega - H$, respectively. Then*

$$(5.10)\qquad \gamma = \delta\alpha, \quad \epsilon = d\beta, \quad \alpha = d\Omega^+, \quad \beta = \delta\Omega^-, \quad d\alpha = \delta\beta = 0.$$

If $\omega \in \mathfrak{P}_2$, then α and β are the plus and minus potentials of $d\omega$ and $\delta\omega$, respectively. We have the following differentiability results on the closure of $\mathfrak{M}$:

(i) *If $\mathfrak{M}$ is of class C_1^1 and $\omega \in \mathfrak{L}_{2\lambda}$, then γ, ϵ, and $H \in \mathfrak{L}_{2\lambda}$ and α, β, Ω^+, and $\Omega^- \in \mathfrak{P}_{2\lambda}$; if also $\omega \in \mathfrak{P}_{2\lambda}$, then γ, ϵ, and $H \in \mathfrak{P}_{2\lambda}$ with γ, δ and H in C_μ^0 in case $\lambda = \rho - 1 + \mu$, $0 < \mu < 1$.*

(ii) *If $\mathfrak{M}$ is of class C_μ^k with $k \geqq 2$* and $0 < \mu < 1$ *and if $\omega \in C_\mu^{k-2}$, then H, γ, and $\epsilon \in C_\mu^{k-2}$ and α, β, Ω^+, and $\Omega^- \in C_\mu^{k-1}$; if, also, $\omega \in C_\mu^{k-1}$, then H, γ, and $\epsilon \in C_\mu^{k-1}$.*

(iii) *If $\mathfrak{M}$ and ω are C^∞ or analytic, so are α, β, γ, ϵ, Ω^+, and Ω^-.*

(iv) *In the case of zero forms, H is a constant, and $\epsilon = 0$.*

If $\omega \in \mathfrak{P}_2$ with $d\omega = 0$ or if $\omega \in \mathfrak{L}_2$ and $\omega = d\eta$ where $\eta \in \mathfrak{P}_2$, then $\gamma = 0$; if $\omega \in \mathfrak{P}_2$ and $\delta\omega = 0$ or if $\omega \in \mathfrak{L}_2$ and $\omega = \delta\eta$ where $\eta \in \mathfrak{P}_2$ then $\epsilon = 0$.

Proof. Suppose, first, that $\omega \in \mathfrak{P}_2$. If we then define α, β, γ, ϵ, Ω^+, and Ω^- by (5.10), the results follow from (5.8) and (5.9). In case ω is merely in $\mathfrak{L}_2$, we use the left sides of (5.9) and approximate, using Theorems 3.3 and 5.2. The regularity results and the last statement follow from the facts that α and β are the respective potentials of $d\omega$ and $\delta\omega$, since $dH = \delta H = 0$, in case $\omega \in \mathfrak{P}_2$. The last results for ω merely in $\mathfrak{L}_2$ follow from Lemma 4.4.

We may now prove a slightly strengthened form of an inequality due to Friedrichs [4]:

THEOREM 5.9. *There is a* $\lambda > 0$ *such that if* $\omega \varepsilon \mathfrak{P}_2 \perp \mathfrak{H}$, *then*

$$D(\omega) \geqq \lambda \| \omega \|^2.$$

Proof. For if $\omega \varepsilon \mathfrak{P}_2 \perp \mathfrak{H}$, then

$$(5.11) \qquad \omega = \gamma + \epsilon, \quad \delta\gamma = d\epsilon = 0, \quad \gamma \varepsilon \mathfrak{P}_2^+, \quad \epsilon \varepsilon \mathfrak{P}_2^-, \quad (\gamma, \epsilon) = 0.$$

Hence, from (5.9) and Theorem 5.2, we see that

$$\begin{aligned} D(\omega) = D(\gamma) + D(\epsilon) &\geqq \lambda^+ \| \gamma \|^2 + \lambda^- \| \epsilon \|^2 \geqq \lambda [2 \| \gamma \|^2 + 2 \| \epsilon \|^2] \\ &\geqq \lambda \| \omega \|^2, \lambda = \min [\lambda^+/2, \lambda^-/2]. \end{aligned}$$

The following theorem completes the analogy with the case for a compact manifold without boundary.

THEOREM 5.10. *If* $\eta \varepsilon \mathfrak{L}_2$ *and is* $\mathfrak{L}_2$*-orthogonal to* $\mathfrak{H}$, *there is a unique form* Ω *in* $\mathfrak{P}_2$ *and* $\mathfrak{L}_2$*-orthogonal to* $\mathfrak{H}$ *such that*

$$(5.12) \qquad (d\Omega, d\zeta) + (\delta\Omega, \delta\zeta) = (\eta, \zeta), \quad \zeta \varepsilon \mathfrak{P}_2.$$

Moreover,

$$(5.13) \qquad d\Omega = d\Omega^+ \text{ and } \delta\Omega = \delta\Omega^-,$$

Ω^+ *and* Ω^- *being the respective plus and minus potentials of* η. *The differentiability properties of* Ω *are the same as those in Theorem* 5.4.

Proof. The proof that Ω exists in $\mathfrak{P}_2$ and is unique is just like that of Theorem 5.3. Obviously $\eta \varepsilon \mathfrak{L}_2 \ominus \mathfrak{H}^+$ and $\mathfrak{L}_2 \ominus \mathfrak{H}^-$ so that its plus and minus potentials exist. Accordingly we have, for example,

$$(5.14) \qquad (d\Omega - d\Omega^+, d\zeta) + (\delta\Omega - \delta\Omega^+, \delta\zeta) = 0 \text{ for all } \zeta \varepsilon \mathfrak{P}_2^+.$$

But, from Theorem 5.6, we may find a $\zeta \varepsilon \mathfrak{P}_2^+$ such that

$$(5.15) \qquad d\zeta = d\Omega - d\Omega^+, \delta\zeta = 0.$$

Using (5.14 and (5.15) and a similar argument for Ω^-, we derive (5.13).

Now, let us consider the decomposition (5.7) for Ω, Ω^+, and Ω^-. Using (5.13) and the last statement in Theorem 5.8 also, we obtain

$$(5.16) \qquad \begin{aligned} &\Omega = \Gamma + E, \quad \Omega^\pm = H^\pm + \Gamma^\pm + E^\pm, \quad \Omega = \Omega^\pm + K^\pm, \quad dK^+ = \delta K^- = 0 \\ &K^+ = H_1^+ + E_1^+, \quad K^- = H_1^- + \Gamma_1^-. \end{aligned}$$

where the Γ's $\varepsilon\, \mathfrak{C}$, the E's $\varepsilon\, \mathfrak{D}$, and the H's $\varepsilon\, \mathfrak{H}$. From (5.16) and the uniqueness of the decomposition, we obtain

$$H^{\pm}+H_1^{\pm}=0, \quad \Gamma+E=\Gamma^{+}+(E^{+}+E_1^{+})=(\Gamma^{-}+\Gamma_1^{-})+E^{-} \tag{5.17}$$
$$\Gamma=\Gamma^{+}, \quad E=E^{-}.$$

The differentiability properties of $d\Omega$ and $\delta\Omega$ follow from (5.13) and Theorem 5.4 and those for Ω follow from (5.16) and (5.17) and Theorems 5.4 and 5.9.

Remark. We cannot conclude the differentiability of Ω directly from (5.12) and Theorem 3.4, since the equations (5.12) are not the same as (3.1) and (3.3) since the boundary integral corresponding to the first term in (4.8) is not necessarily zero in this case.

Definition 5.4. The form Ω in Theorem 5.10 is called the *potential of* η.

Important Remark. All differentiability properties at either interior or boundary points are local; all differentiability results extend immediately to cases where the given hypotheses hold in some coordinate patch, the conclusions then holding in that patch.

6. Boundary value problems. In this section, we derive briefly the results concerning boundary value problems for harmonic forms and fields which have been obtained by other methods by Duff and Spencer [3] and Conner [1]. The differentiability results on the interior have been obtained in part I and stated also in [8]; the corresponding results on the boundary depend on the given boundary values as well as on the differentiability of $\mathfrak{M}$ and are stated below.

The following theorem is seen (from their proofs) to be equivalent to Theorems 3 and 4 of [3] (pp. 150, 151):

THEOREM 6.1. (a) *If ω is any closed form in $\mathfrak{P}_2$, there is a unique harmonic field H such that ($r=$ degree of ω)*

$$\omega=H+d\beta,\ tH=t\omega,\beta,d\beta\,\varepsilon\,\mathfrak{P}_2^{-}(t\beta=td\beta=0), \quad 0\leqq r\leqq n-1.$$

(b) *If ω is any co-closed form in $\mathfrak{P}_2$, there is a unique harmonic field H such that*

$$\omega=H+\delta\alpha,\ nH=n\omega,\alpha,\delta\alpha\,\varepsilon\,\mathfrak{P}_2^{+}\,(n\alpha=n\delta\alpha=0), \quad 1\leqq r\leqq n.$$

In either case, the differentiability results are as follows:

(i) *If $\mathfrak{M}$ is of class C_1^1 and $\omega\,\varepsilon\,\mathfrak{P}_{2\lambda}$, $\lambda=\rho-1+\mu$, $0<\mu<1$, then H is also and ω and $H\,\varepsilon\,C_\mu^0$ at B. If $r=0$, ω and H are constants.*

(ii) *If $\mathfrak{M}$ is of class C_μ^k (C^∞, analytic) and ω is of class C_μ^{k-1} (C^∞, C^ω), $k \geqq 2$, $0 < \mu < 1$, then H is of class C_μ^{k-1} (C^∞, analytic).*

Proof. If $d\omega = 0$, then from Theorems 5.7 and 5.8, we see that the term $\delta\alpha = 0$ in the decomposition which proves (a); (b) is proved similarly. The differentiability results follow from Theorem 5.8.

The next theorem is a refinement of Theorem 2 of [3]:

THEOREM 6.2. *If η is any form in $\mathfrak{P}_2$, there is a form ω in $\mathfrak{P}_2$ such that $t\omega = t\eta$ and $d\omega$ is a harmonic field. If, also, $\eta = \delta\chi$ for some χ in $\mathfrak{P}_2$, then there is a unique ω of the form $\omega = \delta\xi$ with ξ in $\mathfrak{P}_2$ which satisfies the conditions above.*

Proof. Let H^- be the projection of $d\eta$ on $\mathfrak{H}^-$, let α be the minus potential of $d\eta - H^-$, and let

$$\omega = \eta - \gamma, \quad \gamma = \delta\alpha, \quad \epsilon = d\alpha. \tag{6.1}$$

Then γ and ϵ are in $\mathfrak{P}_2^-$ and from (5.3) ($\alpha = \Omega$), we have

$$t\gamma = 0 = t\epsilon, \quad d\gamma + \delta\epsilon = d\eta - H^-, \quad d\omega = d\eta - d\gamma = H^- + \delta\epsilon. \tag{6.2}$$

From (6.2) and the last statement in Theorem 5.8, we see that

$$d\omega \,\varepsilon\, (\mathfrak{H} \oplus \mathfrak{D}) \cap (\mathfrak{H} \oplus \mathfrak{C}) = \mathfrak{H}.$$

Now, suppose $\eta = \delta\chi$ for some χ in $\mathfrak{P}_2$. Then $\omega = \delta(\chi - \alpha)$ from (6.1). Suppose ω_1 also satisfies all these conditions. Then

$$t(\omega - \omega_1) = 0, \quad \omega - \omega_1 = \delta\nu \ (\nu = \chi - \alpha - \xi_1), \quad d(\omega - \omega_1) \,\varepsilon\, \mathfrak{H}.$$

But then, from the definitions of $\mathfrak{C}$ and $\mathfrak{D}$ and Theorem 5.8, we obtain

$$\omega - \omega_1 \,\varepsilon\, \mathfrak{P}_2^-, \ \therefore d(\omega - \omega_1) \,\varepsilon\, \mathfrak{D} \cap \mathfrak{H}, \ \therefore d(\omega - \omega_1) = 0, \ \therefore \omega - \omega_1 \,\varepsilon\, \mathfrak{H} \oplus \mathfrak{D},$$

$$\omega - \omega_1 = \delta\nu \,\varepsilon\, \mathfrak{H} \oplus \mathfrak{C}, \ \therefore \omega - \omega_1 \,\varepsilon\, \mathfrak{H}^-, \ \omega - \omega_1 \,\varepsilon\, \mathfrak{L}_2 - \mathfrak{H}^- \quad \text{(Theorem 5.5(i))},$$

$$\therefore \omega - \omega_1 = 0.$$

We now consider boundary value problems for harmonic forms as distinct from harmonic fields. We begin by defining

$$\mathfrak{P}_{20} = \mathfrak{P}_2^+ \cap \mathfrak{P}_2^-, \quad \mathfrak{H}_0 = \mathfrak{H}^+ \cap \mathfrak{H}^- = \mathfrak{H} \cap \mathfrak{P}_{20}.$$

Then $\mathfrak{P}_{20}$ consists of all $\mathfrak{P}_2$ forms ω with $t\omega = n\omega = 0$. The last part of the following theorem is due to Spencer [11].

THEOREM 6.3. *$\mathfrak{H}_0$ is finite dimensional. If $H \varepsilon \mathfrak{H}_0$, $\mathfrak{H}^+$, or $\mathfrak{H}^-$, then H has the differentiability properties on $\mathfrak{B}$ stated in Theorem* 5 *of* [8]. *If $\mathfrak{M}$ (and $\mathfrak{B}$) is analytic and $H \varepsilon \mathfrak{H}_0$, then $H = 0$.*

Proof. The first statement is a consequence of Theorem 5.1. Obviously

$$(dH, d\zeta) + (\delta H, \delta\zeta) = 0 \tag{6.3}$$

for all ζ in $\mathfrak{P}_2$ and hence in $\mathfrak{P}_2^+$ or $\mathfrak{P}_2^-$. If $H \varepsilon \mathfrak{P}_2^+(\mathfrak{P}_2^-)$, then (6.3) is equivalent to (3.1) and (3.3) so the differentiability results follow.

Now chose any point P on $\mathfrak{B}$ and choose an analytic admissible boundary coordinate system mapping the origin into P and $G_R \cup \sigma_R$ into a neighborhood of P with $g_{ij}(0) = \delta_{ij}$. Since $H \varepsilon \mathfrak{P}_2^+$, (6.3) is equivalent to (3.1) and (3.3) with a, b, b^*, and c analytic and $e = f = 0$. Accordingly H is analytic on $x^n = 0$ and hence H and the coefficients can all be extended analytically across $x^n = 0$. Since $dH = \delta H = 0$ we see that all the $H_{(i)}$ and $H_{(i)x^n} = 0$ on $x^n = 0$. The result follows from the Cauchy-Kowalewsky theorem.

Definition 6.1. A form K is *harmonic* on $\mathfrak{M}$ if and only if K, dK, and $\delta K \varepsilon \mathfrak{P}_2$ on any damain interior to $\mathfrak{M}$ with $\delta dK + d\delta K = 0$ there.

THEOREM 6.4. *If $\omega \varepsilon \mathfrak{P}_2$, there exists a harmonic form K in $\mathfrak{P}_2$ such that $tK = t\omega$, $nK = n\omega$. Any two such solutions differ by a harmonic field in $\mathfrak{H}_0$. The differentiability results for K are the same as those in Theorem* 6.1 *except that in the case of zero-forms, $K \varepsilon C_\mu^k$ if $\omega \varepsilon C_\mu^k$.*

Proof. Write $K = \omega + \eta$ and minimize

$$D(\omega + \eta) = D(\eta) - 2(d\eta, d\omega) + 2(\delta\eta, \delta\omega) + D(\omega)$$

among all η in $\mathfrak{P}_{20} \cap (\mathfrak{L}_2 \ominus \mathfrak{H}_0)$. Then $D(\eta) \geqq \lambda \| \eta \|^2$ so that the minimizing function exists as usual. Then K is easily seen to satisfy (6.3) for all ζ in $\mathfrak{P}_{20}$ so that K is harmonic on the interior of $\mathfrak{M}$ (using the differentiability results of Section 4 and 7 of part I). Since $\eta \varepsilon \mathfrak{P}_{20} \cap (\mathfrak{L}_2 \ominus \mathfrak{H}_0)$, the equivalent equation

$$(d\eta + d\omega, d\zeta) + (\delta\eta + \delta\omega, \delta\zeta) = 0 \tag{6.4}$$

for all $\zeta \varepsilon \mathfrak{P}_{20}$ is of the form (3.1) and (3.3) on boundary coordinate patches. The differentiability results follow from Section 3. In case $K \varepsilon \mathfrak{P}_{20}$, we may set $H = \zeta = K$ in (6.3) and conclude that K is a harmonic field in $\mathfrak{H}_0$.

LEMMA 6.1. *Suppose $f(x)$ is of class $\mathfrak{P}_2$ for $|x| < R$ with support interior to that sphere. Then there is a function $u(x^1, \cdots, x^n, h)$ with*

support in $|x|^2 + h^2 < R^2$, *with* u *and* $\nabla u \in \mathfrak{P}_2$ *there, with* $u(x,0) = 0$, $u_h(x,0) = f(x)$. *If, also* $f(x) \in C_\mu^k$, $0 \leqq \mu \leqq 1$, *then* u *may be taken to be of class* C_μ^{k+1} *there. If, also,* $f \in C^\infty$, *then* u *may be taken to be of class* C^∞. *If, also,* f *is analytic near* x_0, *then* u *may be taken to be analytic in* (x,h) *near* $(x_0, 0)$.

Proof. (By Friedrichs mollifier): Let $k(x)$ be of class C^∞ for all x and have support in $|x| < 1$ with the integral of $k(x) = 1$. Extend $f(x)$ to be zero outside $|x| = R$ and define

$$(6.4) \qquad u_1(x,h) = h \int_{|y|<1} f(x+hy)k(y)\,dy = h^{1-n} \int_{-\infty}^{\infty} f(z)k(z/h - x/h)\,dz.$$

In all cases u_1 is of class C^∞ in (x,h) for $h \neq 0$ and we have

$$(6.5) \qquad \begin{aligned} u_{1x^a}(x,h) &= -\int_{B(0,1)} f(x+h\eta)k_{x^a}(\eta)\,d\eta \\ u_{1h}(x,h) &= -\int_{B(0,1)} f(x+h\eta)[\eta^a k_{x^a}(\eta) + (n-1)k(\eta)]\,d\eta \end{aligned}$$

and $u_{1h}(x,h)$ tends to $f(x)$ as $h \to 0$ for each x for which the Lebesgue derivative of the integral of $f = f(x)$. It is easy to see from (6.5) that u_1 has the desired differentiability properties. Since $f = 0$ for all x on and near $\partial\sigma_R$, we may multiply u_1 by a function $l(h)$ of h alone which is analytic at $h = 0$ with $l(0) = 1$ and $l'(0) = 0$ and C^∞ for all h and zero for $|h| \geqq h_0$ where h_0 is chosen small enough so that $u(x,h) = l(h)u_1(x,h)$ has support in $|x|^2 + h^2 < R^2$.

The following lemma simplifies the statements and proofs of our remaining results on boundary value problems.

LEMMA 6.2. (i) *Suppose* ξ *and* η *are* r *and* $r+1$ *forms on* $\mathfrak{M}$, *respectively, with* ξ, $d\xi$, *and* η *in* $\mathfrak{P}_2$ *along* $\mathfrak{B}$ (*i.e.* $b\xi$, $bd\xi$, $b\eta \in \mathfrak{P}_2$ *along* $\mathfrak{B}$). *Then there exists an* r *form* ω *with* ω *and* $d\omega$ *in* $\mathfrak{P}_2$ *on* $\mathfrak{M}$ *with*

$$(6.6) \qquad n\omega = n\xi,\ nd\omega = n\eta \text{ on } \mathfrak{B}.$$

If ξ, $d\xi$, *and* $\eta \in \mathfrak{P}_{2\lambda}$ *along* $\mathfrak{B}$, ω *may be chosen so that* ω *and* $d\omega \in \mathfrak{P}_{2\lambda}$ *on* $\mathfrak{M}$. *If* $\mathfrak{M}$ *is of class* C_μ^k *and* ξ, $d\xi$, *and* $\eta \in C_\mu^{k-1}$ *along* $\mathfrak{B}$ ($k \geqq 2$, $0 < \mu < 1$), ω *may be chosen so that* ω *and* $d\omega \in C_\mu^{k-1}$. *If* $\mathfrak{M} \in C^\infty$ *and* ξ *and* $\eta \in C^\infty$ *along* $\mathfrak{B}$, ω *may be chosen to be of class* C^∞ *on* $\mathfrak{M}$; *if* $\mathfrak{M}$ *is analytic and* ξ *and* η *are also analytic near a point* P *of* $\mathfrak{B}$, *then* ω *may be chosen to be also analytic on* $\mathfrak{M}$ *near* P. *If* $r = 0$, *the first condition is vacuous and we may choose* $\omega \in C_\mu^k$

if $\mathfrak{M} \varepsilon C_\mu^k$ and $\eta \varepsilon C_\mu^{k-1}$ or to be in $\mathfrak{P}_2$ with all the derivatives of its components being in $\mathfrak{P}_2$ if $\mathfrak{M} \varepsilon C_1^1$. If $r=n$, the second condition is vacuous and $n\omega = b\omega$, $n\xi = b\xi$, and the result is that of Lemma 4.1.

(ii) *The dual of* (i), *with*

$$t\omega = t\xi,\ t\delta\omega = t\eta \text{ on } \mathfrak{B}. \tag{6.7}$$

(iii) *Suppose ξ and η are $r+1$ and $r-1$ forms on $\mathfrak{M}$, respectively, with ξ and η in $\mathfrak{P}_2$ along $\mathfrak{B}$. Then there exist a form ω with ω, $d\omega$, and $\delta\omega$ in $\mathfrak{P}_2$ on $\mathfrak{M}$ such that*

$$nd\omega = n\xi,\ t\delta\omega = t\eta \text{ on } \mathfrak{B}; \tag{6.8}$$

if ξ and $\eta \varepsilon \mathfrak{P}_{2\lambda}$ along $\mathfrak{B}$, ω may be chosen so that ω, $d\omega$, and $\delta\omega \varepsilon \mathfrak{P}_{2\lambda}$ on $\mathfrak{M}$. If $\mathfrak{M}$ is of class C_μ^k ($k \geqq 2, 0 < \mu < 1$) and ξ and $\eta \varepsilon C_\mu^{k-1}$ along $\mathfrak{B}$, then ω may be chosen so that ω, $d\omega$, and $\delta\omega \varepsilon C_\mu^{k-1}$ on $\mathfrak{M}$. If $\mathfrak{M} \varepsilon C^\infty$ and ξ and $\eta \varepsilon C^\infty$ along $\mathfrak{B}$, then ω may be taken in C^∞; if also $\mathfrak{M}$, ξ, and η are analytic near a point P on $\mathfrak{B}$, then ω may be taken analytic on $\mathfrak{M}$ near P. If $r=0$, the condition on $\delta\omega$ is vacuous and ω may be taken to have one additional degree of differentiability (but not $d\omega$); the case $r=n$ is the dual of this (${}^*\omega$ additionally differentiable).

Proof of (i). From the proof of Lemma 4.1, we conclude that there are forms $\xi_1, \cdots, \xi_S$ and $\eta_1, \cdots, \eta_S$ on $\mathfrak{M}$ in which each pair (ξ_s, η_s) has support in some one admissible coordinate patch and each satisfies the differentiability requirements of ξ and η but on $\mathfrak{M}$. Thus our problem is reduced to the case where ξ and η have their stated properties on $\mathfrak{M}$ and support in an admissible boundary coordinate system with domain G_R. We write $\omega = \xi + \omega_1$, $\eta = d\xi + \eta_1$ and seek an ω_1 with the desired differentiability properties satisfying

$$n\omega_1 = 0,\ nd\omega_1 = \eta_1 \text{ on } \mathfrak{B}. \tag{6.9}$$

Actually, we may set $b\omega_1 = 0$ in which case (6.9) is seen using the formula for $d\phi$ (see formula 1.9 of [3]) to reduce to

$$(-1)^r \omega_{1i_1\cdots i_r x^n}(0, x'_n) = \eta_{1i_1\cdots i_r},\ i_r < n,\ \omega_{1i_1\cdots i_{r-1}n}(0, x'_n) = 0. \tag{6.10}$$

From Lemma 6.1, we see that we extend the components to be of class C_μ^k (or etc.) in G_R and to vanish on and near S_R^-. Accordingly ω_1 and $d\omega_1$ have the proper class on $\mathfrak{M}$.

Proof of (ii). We begin by taking the duals of all the forms and then proceeding as in (i).

Proof of (iii). From (i) and (ii) we may find forms ω_1 and ω_2 with ω_1, $d\omega_1$, ω_2, and $\delta\omega_2$ all in C_μ^{k-1} (or etc.) such that

$$n\omega_1 = 0, \ nd\omega_1 = n\xi, \ t\omega_2 = 0, \ t\delta\omega_2 = t\eta \text{ on } \mathfrak{B}.$$

Then from Theorem 5.6(ii), we may find forms ω_3 and ω_4 with

$$n\omega_3 = 0, \ d\omega_3 = d\omega_1, \ \delta\omega_3 = 0, \ t\omega_4 = 0, \ \delta\omega_4 = \delta\omega_2, \ d\omega_4 = 0.$$

The desired $\omega = \omega_3 + \omega_4$. It is seen from Theorems 5.4, 5.5, 5.6, and their proofs and that of hTeorem 5.8, etc., that ω has the desired differentiability properties.

LEMMA 6.3. *Suppose $\omega \varepsilon \mathfrak{L}_2 \ominus \mathfrak{H}$, $\omega \varepsilon \mathfrak{P}_2$ on all domains (with closures) interior to $\mathfrak{M}$, and $D(\omega)$ is finite. Then $\omega \varepsilon \mathfrak{P}_2$ on $\mathfrak{M}$.*

Proof. We may choose a finite covering of $\mathfrak{M}$ by admissible coordinate patches, choose a corresponding partition of unity, and thus express $\omega = \omega_1 + \cdots + \omega_S$, each ω_s having support in some one coordinate patch, being in $\mathfrak{P}_2$ on interior domains, in $\mathfrak{L}_2$ on $\mathfrak{M}$, with $D(\omega_s)$ finite; each ω_s with support interior to $\mathfrak{M} \varepsilon \mathfrak{P}_2$. Now, consider an ω_s with support in a boundary patch with domain G_R; we assume $\omega_s = 0$ elsewhere. Now for each p define $\omega_s{}^p$ on G_R by

$$\omega_{s(i)}{}^p(x^n, x'_n) = \omega_{s(i)}(x^n - h_p, x'_n) \qquad h_p = R\cdot(p+1)^{-1}.$$

If we define $'\omega_s{}^p = \omega_s$ for those with interior support and define $'\omega^p = \omega_1{}^p + \cdots + \omega_S{}^p$ on $\mathfrak{M}$ we see that $'\omega^p$, $d'\omega^p$, and $\delta'\omega^p$ tend strongly in $\mathfrak{L}_2$ on $\mathfrak{M}$ to ω, $d\omega$, and $\delta\omega$, each $'\omega^p$ being in $\mathfrak{P}_2$. If we let ω^p be the projection of $'\omega^p$ on $\mathfrak{L}_2 \ominus \mathfrak{H}$, we see from Theorem 5.8 that we may approximate equally well using $\{\omega^p\}$. But then, from Theorem 5.9 and the fact that $D(\omega^p - \omega) \to 0$, we see that $\|\omega^p\|$ is uniformly bounded. Hence a subsequence, still called ω^p, converges weakly in $\mathfrak{P}_2$ and hence strongly in $\mathfrak{L}_2$ to something which must be ω.

We can now state the results of Conner [1]:

THEOREM 6.5. *In each part of Lemma 6.2, the form ω may be replaced by a harmonic form K with the same differentiability properties. In fact if in* (i) *we merely require that ξ, $d\xi$, and $\eta \varepsilon \mathfrak{P}_2(\mathfrak{P}_{2\lambda})$ on $\mathfrak{M}$, then any such $K \varepsilon \mathfrak{P}_2(\mathfrak{P}_{2\lambda})$ along with dK; analogous results hold in cases* (ii) *and* (iii), *except that in* (iii) *K must be properly chosen, orthogonal to $\mathfrak{H}$ for example. In case* (i) ((ii)), *any two solutions K with K and $dK(\delta K)$ in $\mathfrak{P}_2$ differ by a harmonic field in $\mathfrak{H}^+(\mathfrak{H}^-)$; in case* (iii), *any two solutions K in $\mathfrak{L}_2$ with*

dK and δK in $\mathfrak{L}_2$ differ by a harmonic field. If in (i) ((ii)) $\delta\omega(d\omega)\,\varepsilon\, C_\mu^{k-1}$ *(or etc.), then $\delta K(dK)$ does also.*

Proof of case (i). Let ω_1 be the positive potential of $\delta d\omega$; since $d\omega\,\varepsilon\,\mathfrak{P}_2^+$, $\delta d\omega\,\varepsilon\,\mathfrak{L}_2 \ominus \mathfrak{H}$. Choose ω_2 in $\mathfrak{P}_2^+$ (Theorem 5.6) so that

$$d\omega_2 = 0, \quad \delta\omega_2 = \delta\omega. \tag{6.11}$$

Let us define K by

$$K = \omega - \omega_1 - \omega_2. \tag{6.12}$$

Then we have

$$dK = d\omega - d\omega_1\,\varepsilon\,\mathfrak{P}_2,\ \delta dK = \delta d\omega - \delta d\omega_1 = d\delta\omega_1,\ nK = n\omega,\ ndK = nd\omega. \tag{6.13}$$

Then if $\zeta\,\varepsilon\,\mathfrak{P}_{20}$, we have

$$\begin{aligned}(dK, d\zeta) + (\delta K, \delta\zeta) &= (\delta dK, \zeta) + (\delta K, \delta\zeta)\\ &= (d\delta\omega_1, \zeta) + (\delta K, \delta\zeta) = (\delta K + \delta\omega_1, \delta\zeta) = (\delta\omega - \delta\omega_2, \delta\zeta) = 0\end{aligned}$$

using (6.11). Hence K is harmonic from Sections 4 and 7, part I. If $\delta\omega$ were already known to be in $C_\mu^{k-1}(\mathfrak{P}_2$, etc.), we would simply have defined K by

$$K = \omega - \Omega \tag{6.14}$$

where Ω is the plus potential of $\Delta\omega$.

Now, in the case of two solutions, the *difference* K' would be a harmonic form with K' and dK' in $\mathfrak{P}_2$ and

$$nK' = ndK' = 0. \tag{6.15}$$

Now $\delta K'\,\varepsilon\,\mathfrak{P}_2$ on interior domains is harmonic there (Sections 4 and 7, part I) and is in $\mathfrak{L}_2$ on $\mathfrak{M}$. Moreover

$$d\delta K' = -\delta dK'\,\varepsilon\,\mathfrak{L}_2, \quad \delta\delta K' = 0.$$

Hence $D(\delta K')$ is finite and, by Definition 5.3, $\delta K'\,\varepsilon\,\mathfrak{C} \subset \mathfrak{L}_2 \ominus \mathfrak{H}$. Hence $\delta K'$ is also in $\mathfrak{P}_2$ by Lemma 6.3. But then

$$\begin{aligned}(dK', dK') + (\delta K', \delta K') &= (K', \Delta K')\\ &\quad + \int_{\mathfrak{B}} K' \cap {}^*dK' - \delta K' \cap {}^*K' = 0,\end{aligned} \tag{6.16}$$

so that $K'\,\varepsilon\,\mathfrak{H}^+$.

Proof of case (ii). Dual to case (i).

Proof of case (iii). In this case we define K by (6.14) and note that it satisfies the conditions if we choose Ω as the potential of

$$\Delta\omega = \delta(d\omega) + d(\delta\omega) \varepsilon \mathfrak{C} \oplus \mathfrak{D}$$

(Definition 5.3, Theorem 5.10).

Suppose now that we have two solutions with the stated properties. Then the projection K' of the difference on $\mathfrak{C} \oplus \mathfrak{D}$ is harmonic and in $\mathfrak{L}_2 \ominus \mathfrak{H}$ with dK' and $\delta K'$ in $\mathfrak{P}_2$ with $ndK' = t\delta K' = 0$. Then (6.16) holds so that $K' \varepsilon \mathfrak{H}$. Hence $K' = 0$.

Remarks. We conclude with remarks about a result of Spencer on what he called a bounded manifold [10]. An open manifold $\mathfrak{M}$ is *bounded* if and only if there is a compact manifold $\mathfrak{M}_1$, with or without boundary, of which $\mathfrak{M}$ is an open subdomain with closure interior to $\mathfrak{M}_1$; as usual, we assume $\mathfrak{M}_1$ of class $C_1{}^1$ at least. We may define the subspace $\mathfrak{P}_{20}$ of $\mathfrak{P}_2$ forms on $\mathfrak{M}$ as the closure in $\mathfrak{P}_2$ on $\mathfrak{M}_1$ of the $\mathfrak{P}_2$ (or Lipschitz) forms with support interior to $\mathfrak{M}$; any form in $\mathfrak{P}_{20}$ is in $\mathfrak{P}_2$ on $\mathfrak{M}_1$ if it is defined to be zero on and outside $\partial\mathfrak{M}$. Thus, if $\mathfrak{N}$ is any finite covering of $\mathfrak{M}_1$, we have

$$D(\omega) \geqq k \| \omega \|^2 - l(\omega, \omega)$$

for all ω in $\mathfrak{P}_{20}$ on $\mathfrak{M}$ the norm corresponding to $\mathfrak{N}$. Clearly $D(\omega)$ is still lower-semicontinuous with respect to weak convergence in $\mathfrak{P}_{20}$ and weak convergence in $\mathfrak{P}_{20}$ implies strong convergence in $\mathfrak{L}_2$. Hence we may generalize Lemma 5.1 and Theorems 5.1 to 5.3 immediately and hence also Theorems 6.3 and 6.4 except for the differentiability on the boundary; the conditions $tK = t\omega$ and $nK = n\omega$ of Theorem 6.4 must be changed to $K - \omega \varepsilon \mathfrak{P}_{20}$. In Theorem 6.3, if $\mathfrak{M}_1$ is analytic, we see that any harmonic field in $\mathfrak{P}_{20}$ is also one on $\mathfrak{M}_1$ if extended as above; since $\mathfrak{M} \cup \partial\mathfrak{M}$ is interior to $\mathfrak{M}_1$, it follows that any such field is zero. Moreover if K is a harmonic *form* [defined as in Definition 6.1] in $\mathfrak{P}_{20}$, we see (by first considering ζ's with support interior to $\mathfrak{M}$) that $(dK, d\zeta) + (\delta K, \delta\zeta) = 0$ for all ζ in $\mathfrak{P}_{20}$; by taking $\zeta = K$, we see that K is a harmonic field.

UNIVERSITY OF CALIFORNIA AND
THE INSTITUTE FOR ADVANCED STUDY.

REFERENCES.

[1] P. E. Conner, "The Green's and Neumann's problems for differential forms on Riemannian manifolds," *Proceedings of the National Academy of Science, U. S. A.*, vol. 40 (1954), pp. 1151-1155.

[2] G. de Rham and K. Kodaira, "Harmonic Integrals," Mimeographed Notes, Institute for Advanced Study, 1950.

[3] G. F. D. Duff and D. C. Spencer, "Harmonic tensors on Riemannian manifolds with boundary," *Annals of Mathematics*, vol. 56 (1952), pp. 128-156.

[4] K. O. Friedrichs, "On differential forms on Riemannian manifolds," *Communications on Pure and Applied Mathematics*, to appear.

[5] M. Gaffney, "The harmonic operator for exterior differential forms," *Proceedings of the National Academy of Science, U. S. A.*, vol. 37 (1951), pp. 48-50.

[6] K. Kodaira, "Harmonic fields in Riemannian manifolds," *Annals of Mathematics*, vol. 50 (1949), pp. 587-665.

[7] C. B. Morrey, Jr., "Second order elliptic systems of differential equations," *Annals of Mathematics, Studies*, no. 33, Princeton Univ. Press, 1954.

[8] ——— and James Eells, Jr., "A variational method in the theory of harmonic integrals," *Proseedings of the National Academy of Science, U. S. A.*, to appear.

[9] ——— and James Eells, Jr., "A variational method in the theory of harmonic integrals I," *Annals of Mathematics*, to appear.

[10] D. C. Spencer, "Dirichlet's principle on manifolds," *Studies in Mathematics and Mechanics presented to Richard von Mises, Academic Press* (1954), pp. 127-134.

[11] ———, "Almost Complex Structure," Mimeographed Notes, Princeton University, 1953-54.

FIBRE SYSTEMS OF JACOBIAN VARIETIES.* [1]

By Jun-ichi Igusa.

The method of Picard and Poincaré in studying an algebraic surface V consists in applying the theory of curves to the curves of an auxililary linear pencil $\{C_u\}$ on V. The same method was employed later by Lefschetz in his theory of algebraic surfaces (cf. [19], Chapters 6-7). If we examine closely their method, we can see that they considered implicitly a certain variety $\mathcal{J}$ which the author proposes to call the *Néron variety of V associated with* $\{C_u\}$. In fact Néron was the first who considered $\mathcal{J}$ explicitly in his algebraic proof of the theorem of the base [12]. The variety $\mathcal{J}$ is the graph of the correspondence $u \rightarrow J_u$ between u and the Jacobian variety J_u of C_u. Here, we restrict our attention to such linear pencils whose members are all irreducible and whose general members are nonsingular. If V does not carry any multiple curve, we can always find such a linear pencil. Also, if V is nonsingular, we can assume that singular members of the pencil are curves with ordinary double points. Now, the main part of the paper is devoted to determining the "degenerate fibres" of the fibre system $\{u \times J_u\}$ on $\mathcal{J}$ at those finite number of values of u where C_u become singular. We note that in Néron's case such degenerate fibres are not considered explicitly. The same thing can be said about Chow's investigations on Abelian varieties over function fields [4]. However, in some problems in algebraic geometry it becomes necessary to consider those degenerate fibres and also the behavior of fibres along the degenerate fibres. We shall show that the degenerate fibres are certain completions of the generalized Jacobian varieties of the singular curves in the sense of Rosenlicht [15]. The singular locus of $\mathcal{J}$ is contained in the union of singular loci of degenerate fibres. Also we can define in a natural way a birational map ϕ from V into $\mathcal{J}$, and we can show that ϕ gives isomorphisms of the Albanese varieties and of the spaces of linear differential forms of the first kind of V and $\mathcal{J}$. This result has already been applied to show that the dimension of the Albanese variety of an arbitrary

* Received December 28, 1954; revised September 30, 1955.

[1] This work was supported by a research project at Harvard University, sponsored by the Office of Ordnance Research, United States Army, under Contract DA-19-020-ORD-3100.

variety is not greater than the dimension of the space of linear differential forms of the first kind on the variety [6].

In the course of the preparation of the paper, the author had precious discussions with and valuable suggestions from Professor Zariski, to whom the author wishes to express his deepest appreciation.

1. Existence of general linear pencils.

1. Let V be an arbitrary algebraic variety in a projective space P^n. We shall denote by $\mathfrak{K}$ the universal domain of our algebraic geometry.[2] Let H be a hypersurface of order m in P^n not containing V. Then the intersection product $V \cdot H$ is defined, and $V \cdot H$ is a P^n-cycle carried by V. We shall denote by $\mathfrak{L}_m$ the totality of such cycles. Now the system of hypersurfaces of order m in P^n defines a biregular rational map ϕ of P^n into a projective space of dimension $C_n^{m+n}-1$. Consequently V is transformed biregularly to an algebraic variety V^* in this projective space. We shall denote by P^N the smallest subspace containing V^*. If L is a hyperplane in P^N and if T is the graph of $\mathrm{Tr}_V\,\phi$, then $\mathrm{pr}_V[T \cdot (V \times L)]$ is a member of $\mathfrak{L}_m$. Moreover this operation defines a one-to-one correspondence between the system of hyperplanes in P^N and $\mathfrak{L}_m$. The system of hyperplanes in P^N can be identified with the dual space P'^N of P^N. If P'^r is a subspace of P'^N and if X_0 is a P^n-cycle of dimension $\dim(V)-1$ carried by V, then the set of P^n-cycles of the form $X+X_0$, where X vary in P'^r, is called a *linear system of dimension r on V*. In particular $\mathfrak{L}_m$ itself is a linear system of dimension N on V. In the following V is either an algebraic surface or an algebraic curve.

Now let V be an arbitrary algebraic surface. We know in general that of all fields in $\mathfrak{K}$ which define V, there is one smallest one k_0. We shall denote its algebraic closure by k. Since the rational map ϕ is defined over the prime field, k_0 is again the smallest field of definition of V^*. Therefore the projective space P^N is defined over k_0. A positive cycle is called a reducible variety if the sum of the coefficients of its components is at least equal to 2. A normally algebraic bunch over a field will be called, after Zariski, a closed set over this field. We shall now prove the following theorem.[3]

---

[2] We shall use the results and terminology of Weil's book [16] mostly without quoting.

[3] A slightly less general result was proved by Néron and Samuel [13]. We shall use their idea in our proof.

THEOREM 1. *The set $\mathfrak{C}_1$ of hyperplanes in P^N, which have reducible curves as intersection products with V^*, is a closed set over k_0 of dimension at most $N-2$ for $m \geqq 3$. The only exceptional surface for $m=2$ is a subspace P^2 of P^n.*

Proof. The assertion that $\mathfrak{C}_1$ is a closed set over k_0 can be proved by the standard method [1], hence we can omit it here. We shall first treat the case where V is a subspace of P^n. Let $G(X)$ and $H(X)$ be homogeneous polynomials in three letters X_0, X_1, X_2 of orders α and β respectively such that $\alpha+\beta=m$. Put $F(X)=G(X)H(X)$. The set of all $F(X)$ of this form is an algebraic variety of co-dimension $C_2^{m+2}-(C_2^{\alpha+2}+C_2^{\beta+2})+1 = \alpha(m-\alpha)$ in the projective space of homogeneous polynomials of order m. If we remark that the minimal value of $\alpha(m-\alpha)$ for $1 \leqq \alpha \leqq m-1$ is $m-1$, the first assertion and the trivial part of the second assertion for $V=P^2$ follows immediately. We shall exclude this case in the following. We fix one m by $m \geqq 2$. Let R be a component of $\mathfrak{C}_1$. Then R is defined over $\bar{k}$, and we have $\dim(R) \leqq N-1$ by a theorem of Bertini [21, 9]. We shall derive a contradiction from the assumption $\dim(R)=N-1$. Let u be a generic point of R over k and let L_u be the corresponding hyperplane in P^N. Then $V^* \cdot L_u$ is a rational P^N-cycle over $k(u)$, hence it determines a rational positive cycle X of the product $P^N \times R$ over k, every component of which has the projection R on R and such that $X(u)=V^* \cdot L_u$. Let $x^* \times u$ be a generic point of some component of X over k. Then x^* is a generic point of V^* over k. Otherwise x^* has a locus C^* over k, which is contained in every hyperplane of R. However, this is impossible by the following reason: Let C be the corresponding curve on V. Then we have $\operatorname{ord}(C^*)=m \cdot \operatorname{ord}(C) \geqq 2$. Therefore C^* contains three points which are not colinear. Hence the hypersurface R is contained in the intersection of dual hyperplanes of these three points; a contradiction. Let C^* be a component of $V^* \cdot L_u$ containing x^*. Then C^* appears with coefficient one in $V^* \cdot L_u$, and it is the only component of $V^* \cdot L_u$ passing through x^*. Otherwise, by a criterion of multiplicity one [16, p. 141] L_u contains the tangent plane of V^* at x^*. The totality of hyperplanes in P^N having this property is a subspace T of P'^N of dimension $N-3$ defined over $k_0(x^*)$. On the other hand if K is the algebraic closure of $k_0(x^*)$, the point u has a locus S of dimension $N-2$ over K, and every hyperplane of S contains the tangent plane of V^* at x^*, i.e., S is contained in T; a contradiction. Since C^* appears with coefficient one in $X(u)$, it is defined over a separably algebraic extension of $k(u)$. Since C^* is also the only component of $X(u)$ passing through x^*, we conclude that C^* is defined

over $k(u, x^*)$, hence over $K(u)$. Therefore, by assumption $V^* \cdot L_u - C^*$ is a strictly positive and rational P^N-cycle over $K(u)$. Hence it determines a strictly positive and rational cycle Y of the product $P^N \times S$ over K, every component of which has the projection S on S and such that $Y(u) = V^* \cdot L_u - C^*$. Since S is of dimension $N-2$, our previous argument shows again that every component of Y has the projection V^* on V^*. Let Z be a component of Y, and let T' be the projection on S of a component of the intersection $Z \cap x^* \times S$. Since this intersection is nonempty, T' exists actually and it is a subvariety of S of dimension at least $N-3$ defined over K. From the same criterion of multiplicity one as before, we conclude that T' is contained in T. Since the dimension of T' is at least equal to the dimension of T, we get $T = T'$; hence in particular T is contained in $\mathfrak{C}_1$.

Now let $\bar{x}^*$ be a generic point of V^* over $k_0(x^*)$. Let x and $\bar{x}$ be the points of V corresponding to x^* and $\bar{x}^*$. We may assume that they have representatives $(1, x_1, \cdots, x_n)$ and $(1, \bar{x}_1, \cdots, \bar{x}_n)$ respectively. Let $X_0, X_1, \cdots, X_n$ be the letters to describe equations in P^n. Then m-fold products of the form

$$(X_i - x_i X_0)(X_j - x_j X_0) \cdots (X_k - x_k X_0)$$

for $1 \leqq i \leqq j \leqq \cdots \leqq k \leqq n$ correspond to hyperplanes in P^N, and, by the criterion of multiplicity one, all of them contain the tangent plane of V^* at x^*. They determine a subspace of T. We may assume $\bar{x}_1 \neq x_1$. Put $y_i = \bar{x}_i - x_i/\bar{x}_1 - x_1$ for $i = 1, \cdots, n$. We shall show that $k_0(x)(\bar{x})$ is separably algebraic over $k_0(x)(y)$. Otherwise, there exists at least one nontrivial derivation D of $k_0(x)(\bar{x})$ over $k_0(x)(y)$. Let $F_j(X)$ be a polynomial in $X_1, \cdots, X_n$ with coefficients in k_0 such that $F_j(\bar{x}) = 0$. We then get $y_i D\bar{x}_1 - D\bar{x}_i = 0$ $(2 \leqq i \leqq n)$, $\sum_{i=1}^{n} \partial F_j/\partial \bar{x}_i \cdot D\bar{x}_i = 0$. Therefore, the determinant of these n linear equations must be zero. Hence we get $\sum_{i=1}^{n} \partial F_j/\partial \bar{x}_i \cdot (x_i - \bar{x}_i) = 0$ for every $F_j(X)$. However, since x is a generic point of V over $k_0(\bar{x})$, this shows that V is contained in, hence coincides with the tangent plane of V at $\bar{x}$. This is the case we have excluded in the beginning. Thus the linear system on V which is determined by T is not "composite with an algebraic pencil," hence its general member must be irreducible by the theorem of Bertini [21, 9]. This contradicts to the conclusion we have arrived before, i.e., to the fact that T is contained in $\mathfrak{C}_1$.

2. The following three lemmas hold also for an arbitrary algebraic surface V:

LEMMA 1. *Let x^* be a simple point of V^*. Then we can find a hyperplane L in P^N containing the tangent plane of V^* at x^* such that $V^* \cdot L$ consists of two nonsingular curves with distinct tangents locally at x^* for $m \geqq 2$.*

Proof. Let x be the point of V which corresponds to x^*. Let L^{n-1} and H^{n-1} be independent generic hyperplane and hypersurface of order $m-1$ over $k_0(x)$ passing both through x. Then $V \cdot L^{n-1}$ and $V \cdot H^{n-1}$ are irreducible curves at least locally at x which are transversal to each other at x on V. Let L' be the hyperplane in P^N which corresponds to $L^{n-1} + H^{n-1}$. Then $V^* \cdot L'$ corresponds biregularly to $V \cdot (L^{n-1} + H^{n-1})$. Therefore $V^* \cdot L'$ consists of two nonsingular curves with distinct tangents locally at x^*.

LEMMA 2. *Let x^* be a simple point of V^*. Then the tangent plane of V^* at x^* meets V^* only at x^* for $m \geqq 2$.*

Proof. We can use the same notations as in the above proof. Suppose that $\bar{x}^*$ is a point of V^* distinct from x^*. Let $\bar{x}$ be the corresponding point of V. If L^{n-1} and H^{n-1} are taken to be generic over $k_0(x, \bar{x})$, then L' does not pass through $\bar{x}^*$. Therefore the tangent plane of V^* at x^* can not contain $\bar{x}^*$.

LEMMA 3. *Let x^* and $\bar{x}^*$ be two distinct simple points of V^*. Then the tangent planes of V^* at x^* and at $\bar{x}^*$ do not intersect with each other for $m \geqq 3$.*

Proof. Let x and $\bar{x}$ be the points of V which correspond to x^* and $\bar{x}^*$. Let H_1^{n-1} and H_2^{n-1} be independent generic hypersurfaces of order $m-2$ over $k_0(x, \bar{x})$ passing both through x. Also let L_1^{n-1}, $L_1'^{n-1}$, L_2^{n-1} and $L_2'^{n-1}$ be four hyperplanes passing through $\bar{x}$, but not through x. Let L_i be the hyperplanes in P^N which correspond to $L_i^{n-1} + L_i'^{n-1} + H_i^{n-1}$ for $i = 1, 2$. Then both L_1 and L_2 contain the tangent plane of V^* at $\bar{x}^*$, while $V^* \cdot L_1$ and $V^* \cdot L_2$ are transversal to each other at x^* on V^*. Therefore $L_1 \cdot L_2$ is a subspace of P^N of dimension $N-2$ which contains the tangent plane of V^* at $\bar{x}^*$ and which is transversal to V^* at x^*. The existence of such a subspace implies that the tangent planes of V^* at x^* and at $\bar{x}^*$ are disjoint.

3. We shall now assume that V has only "negligible singularities," i.e., that V does not contain any multiple curve. We can then add the following three supplements to Theorem 1:

SUPPLEMENT 1. *The set $\mathfrak{F}$ of hyperplanes in P^N, which are not trans-*

versal to V^* at one point of V^* at least, is a closed set over k_0 of dimension at most $N-1$ for $m \geqq 1$.

Proof. First of all the singular locus of V^* is a closed set over k_0 consisting of a finite number of points. The dual hyperplanes of these points form a closed set over k_0 in P'^N. On the other hand, let x^* be a generic point of V^* over k_0 and let L be a generic hyperplane of P^N over $k_0(x^*)$ containing the tangent plane of V^* at x^*. If u is the dual point of L, then u has a locus over k_0 of dimension at most $N-1$. Let L' be a specialization of L over k_0 and let x'^* be a specialization of x^* over this specialization. If x'^* is a multiple point of V^*, then L' is a member of the dual hyperplane of x'^*. If x'^* is a simple point of V^*, then L' contains the tangent plane of V^* at x'^*. This follows from the fact that the tangent plane of V^* at x'^* is the unique specialization of the tangent plane of V^* at x^* over the specialization $x^* \rightarrow x'^*$ with reference to k_0. Conversely let L' be a hyperplane in P^N containing a tangent plane of V^* at a simple point x'^* of V^*. Then L' is a specialization of L over k_0, as we can show by the following general argument: The totality of hyperplanes in P^N containing the tangent plane of V^* at x^* is a subspace T of P'^N of dimension $N-3$ defined over $k_0(x^*)$. Similarly a subspace T' of P'^N of dimension $N-3$ is attached to x'^*. Since the specialization T'' of T over the specialization $x^* \rightarrow x'^*$ with reference to k_0 is carried by T' and since T' and T'' are linear spaces of the same dimension, we have $T'' = T'$. In particular a generic member of T' over $k_0(x'^*)$ is a specialization of L over the specialization $x^* \rightarrow x'^*$ with reference to k_0. Since L' is a member of T' and since T' is defined over $k_0(x'^*)$, we see that L' is a specialization of that generic member over $k_0(x'^*)$. Therefore L' is a specialization of L over the specialization $x^* \rightarrow x'^*$ with reference to k_0. Since a hyperplane L' of P^N is not transversal to V^* at a point x'^* of their intersection if and only if either x'^* is multiple on V^* or x'^* is simple on V^* and L' contains the tangent plane of V^* at x'^*, our assertion is proved.

SUPPLEMENT 2. *The set of hyperplanes in P^N, which are not transversal to V^* at two points of V^* at least, is contained in a closed set $\mathfrak{E}_2$ over k_0 of dimension at most $N-2$ for $m \geqq 3$.*

Proof. Since the set of singular points of V^* is closed over k_0, the set of hyperplanes in P^N containing at least two of them is a finite set of subspaces of dimension $N-2$ in P'^N, which is again closed over k_0. In the next place let a^* be one of the multiple points of V^*, and let x^* be a generic

point of V^* over k_0, hence also over $k_0(a^*)$.. Then we can speak of a generic hyperplane L of P^N over $k_0(x^*, a^*)$ containing the tangent plane of V^* at x^* and passing through a^*. The dual point of L has a locus over $k_0(a^*)$, and the dimension of this locus is at most $N-2$ by Lemma 2. The union of such varieties for all a^* is closed over k_0. Moreover, if L' is a hyperplane in P^N containing a tangent plane of V^* at a simple point x'^* of V^* and passing through a multiple point a^* of V^*, then L' is a specialization of a generic L corresponding to the same a^* over the specialization $x^* \to x'^*$ with reference to $k_0(a^*)$. This can be proved exactly in the same way as in the proof of Supplement 1 using Lemma 2. Finally let x^* and $\bar{x}^*$ be independent generic points of V^* over k_0. Then we can speak of a generic hyperplane L of P^N over $k_0(x^*, \bar{x}^*)$ containing the tangent planes of V^* at x^* and at $\bar{x}^*$. The dual point of L has a locus over k_0, and the dimension of this locus is at most $N-2$ by Lemma 3. Moreover, if L' is a hyperplane in P^N containing tangent planes of V^* at two simple points x'^* and $\bar{x}'^*$, then L' is a specialization of L over the specialization $(x^*, \bar{x}^*) \to (x'^*, \bar{x}'^*)$ with reference to k_0. This can be proved in the same way as in the proof of Supplement 1 using Lemma 3. We can take as $\mathfrak{C}_2$ the union of the above three types of closed sets over k_0.

Our final supplement is more delicate. Let x^* be a simple point of V^*, and let C^* be a positive divisor of V^* passing through x^*. Then we can find an element f of the local ring of V^* at x^* such that its divisor (f) represents C^* locally at x^*. We shall assume that f is contained in the square, but not in the cube of the maximal ideal of the local ring. Then the residue class of f with respect to the cube of the maximal ideal is a quadratic form in the Zariski tangent space of V^* at x^*, and this quadratic form is uniquely determined up to a constant factor by C^*. If its discriminant is not zero, then x^* is called an *ordinary double point* of C^*.

SUPPLEMENT 3. *The set of hyperplanes in P^N, which intersect with V^* along curves with non-ordinary multiple points at simple points of V^*, is contained in a closed set $\mathfrak{C}_3$ over k_0 of dimension at most $N-2$ for $m \geqq 3$.*

Proof. Let x^* be a generic point of V^* over k_0, and let L be a generic hyperplane over $k_0(x^*)$ containing the tangent plane of V^* at x^*. Then x^* is the only point of V^* at which L is not transversal to V^*. Otherwise, from Lemmas 2, 3 we conclude that the dual point u of L is of dimension at most $N-4$ over $k_0(x^*)$. However, this dimension is actually $N-3$, and this is a contradiction. In particular u is not contained in any dual

hyperplane of a multiple point of V^*. Also the point x^* has no other specialization than itself over $k_0(u)$, hence x^* is purely inseparable over $k_0(u)$. Hence the locus U of u over k_0 is of dimension exactly equal to $N-1$. A nonempty subset of U is called to be open over k_0, if its complement is closed over k_0. Let U_0 be an affine representative of U in which u has a representative (u). We note that U_0 is an open set over k_0. Let U_1 be the subset of U which is obtained from U_0 by subtracting the singular locus, the union of dual hyperplanes of multiple points of V^* and the set $\mathfrak{E}_2$ which we introduced in Supplement 2. Then, by our previous remarks U_1 is nonempty, hence it is open over k_0. Moreover, if (u') is a point of U_1, the corresponding hyperplane in P^N is not transversal to V^* at exactly one simple point x'^* of V^*. Thus we have a single-valued map $(u')\rightarrow x'^*$ of U_1 into V^*. Let V_0^* be any, but fixed affine representative of V^*. Let (x^*) be the corresponding representative of x^*. We may assume that (x_1^*, x_2^*) are separating variables of $k_0(x^*)$ over k_0. Then we can find $N-2$ polynomials $F_j(X)$ in the defining ideal over k_0 of V_0^* such that $\det(\partial F_j/\partial x_i^*)_{i\neq 1,2}$ is not zero. Since x^* is purely inseparable over $k_0(u)$, a certain power of this determinant is an element $\phi(u)$ of $k_0(u)$. Here ϕ is a function on U_1 defined over k_0. In the same way we can associate to each x_i^* a function ϕ_i on U_1 defined over k_0. Let U_2 be the complement of U_1 of the poles of ϕ_i and the zeros of ϕ. Then U_2 is open over k_0. Moreover, if (u') is a point of U_2, the corresponding point x'^* on V^* has a representative in V_0^*, and (x_1^*, x_2^*) form local coordinates of V^* at x'^*. We shall denote by $\partial/\partial x_\alpha^*$ the derivations of $k_0(x^*)$ over $k_0(x_\beta^*)$ normalized by $\partial x_\alpha^*/\partial x_\alpha^*=1$ for $\alpha,\beta=1,2$ $(\alpha\neq\beta)$. Then $\det(\sum_i u_i\partial^2 x_i^*/\partial x_\alpha^*\partial x_\beta^*)$ is purely inseparable over $k_0(u)$. Therefore, its certain power determines a function ψ on U_2 over k_0. We note that ψ is regular on U_2 (cf. [8]), and ψ is not identically zero by Lemma 1. Therefore, if we denote by $\mathfrak{E}_3$ the union of the boundary of U_2 and the zeros of ψ in U_2, then $\mathfrak{E}_3$ is a closed set over k_0 of dimension at most $N-2$. It follows from the construction that $\mathfrak{E}_3$ satisfies our demands.

In the following we fix one m by $m\geqq 3$. If u is a point of the dual space P'^N of P^N, and if L_u is the corresponding hyperplane in P^N, then the biregular transform C_u on V of $V^*\cdot L_u$ is a member of $\mathfrak{L}_m$, and C_u is rational over $k_0(u)$. If we define $\mathfrak{E}$ as the union of $\mathfrak{E}_i$ for $i=1,2,3$, then we can summarize our results in the following way:

CONCLUSION. *There exist two closed sets* $\mathfrak{F}$ *and* $\mathfrak{E}$ *over* k_0 *in* P'^N *of dimensions at most* $N-1$ *and* $N-2$ *respectively such that* (i) *if* u *is not*

contained in $\mathfrak{F}$, the corresponding C_u is an irreducible[4] *nonsingular curve*; (ii) *if u' is a point of $\mathfrak{F}-\mathfrak{E}$, then $C_{u'}$ is still irreducible and has one and only one multiple point*; (iii) *if this multiple point is simple on V, then it is an ordinary double point of C_u.*

Now, consider the Grassmann variety G of straight lines in P'^N. The variety G has a generic point over k_0, and this point generates a purely transcendental extension over k_0. Moreover the totality of straight lines in P'^N, which either lie on $\mathfrak{F}$ or intersect with $\mathfrak{E}$, is a closed subset of G over k_0 of co-dimension at least equal to one. On the other hand to each straight line D in P'^N corresponds a subspace L of P^N of dimension $N-2$. If D is generic over k_0, the intersection product $V^* \cdot L$ is defined and consists of, say β, distinct generic points of V^* over k_0. If we associate the Chow point of $V^* \cdot L$ to D, we get a rational map over k_0 of G into the β-fold symmetric product $V^*(\beta)$ of V^*. The points of $V^*(\beta)$, which correspond to β non-distinct points of V^*, form a closed set over k_0 of co-dimension one. The inverse image of this closed set in G is again a closed set over k_0 of co-dimension one. We take an algebraic point over k_0 in G which is contained neither in this set nor in the closed set we introduced before. If k_0 is an infinite field, even a rational point over k_0 can be found under the above condition. It is an open question whether we can do the same thing, by taking a larger m if necessary, when k_0 is finite. Let D be the straight line in P'^N which corresponds to the point under consideration. Then we get a linear pencil $\{C_u\}$ with the parameter straight line D such that the parametrization, i.e., the correspondence $u \rightarrow C_u$ is defined over the algebraic closure k of k_0. Since D does not intersect with $\mathfrak{E}$, (i) *every C_u is irreducible.* Since D is not contained in $\mathfrak{F}$, (ii) *there is a finite number of points, say $a_1, \cdots, a_\alpha$, on D such that C_u become singular only at these points.* (iii) *Each C_{a_i} has one and only one multiple point, and if this multiple point is simple on V, it is an ordinary double point of C_{a_i} for $i=1, \cdots, \alpha$.* Also since we have avoided another kind of exceptional straight lines, by the criterion of multiplicity one, (iv) *the base points $A_1, \cdots, A_\beta$ of the pencil are simple on V, and any two members of the pencil are transversal to each other on V at $A_1, \cdots, A_\beta$.* A linear pencil $\{C_u\}$ with the above four properties will be called a *general linear pencil* on V.

---

[4] It is clear that C_u is a disjoint sum of nonsingular curves. However C_u must be connected by the principle of degeneration. We can avoid this deep principle simply by taking the union of $\mathfrak{F}$ and $\mathfrak{E}$ as our $\mathfrak{F}$.

2. Construction of generalized Jacobian varieties by Chow's method.

4. Let C be an arbitrary irreducible curve in a projective space P^n and let $\mathfrak{o}$ be the intersection of local rings of C at its multiple points. Then $\mathfrak{o}$ determines C uniquely up to biregular transformations. We shall denote by K_0 the smallest field of definition of C. Also, if $\mathfrak{q}$ is a positive 0-cycle in P^n, its Chow point will be denoted by $(\mathfrak{q})$. If K is a field, we write $K(\mathfrak{q})$ instead of $K((\mathfrak{q}))$. On the other hand, a divisor of C means always a divisor of C in the sense of Weil, i.e., every component of a divisor shall be a simple point of C. If f is a function on C, we shall consider its *divisor in the sense of valuation theory.* This is not a good terminology; it means the regular image on C of the divisor of the function induced by f on the normalization of C.

Now, if $\mathfrak{p}$ is an arbitrary divisor of C and if $\mathfrak{q}$ is a 0-cycle in P^n carried by C, we write $\mathfrak{p}\rightarrow\mathfrak{q}$ whenever there exists an element f of $\mathfrak{o}$ such that $\mathfrak{q}-\mathfrak{p}$ is the divisor of f in the sense of valuation theory [14]. Here f does not vanish at any multiple point of C if and only if $\mathfrak{q}$ is also a divisor of C. Therefore the relation $\mathfrak{p}\rightarrow\mathfrak{q}$ can be reversed if and only if $\mathfrak{q}$ is a divisor of C. In other words, if we restrict to divisors, the arrow relation is an equivalence relation. If $\mathfrak{p}$ is a divisor of C, we shall denote by $|\mathfrak{p}|$ the set of all positive 0-cycles $\{\mathfrak{q}\}$ in P^n satisfying $\mathfrak{p}\rightarrow\mathfrak{q}$, and we call $|\mathfrak{p}|$ the *complete linear system on C determined by $\mathfrak{p}$*. We note that every member of $|\mathfrak{p}|$ has the same degree, and we call it the degree of $|\mathfrak{p}|$. Here we must show that $|\mathfrak{p}|$ is actually a linear system on C. This can be done by using the following lemma due to Zariski [20]: *Let V be an arbitrary variety in P^n, and let $t_0,\cdots,t_n$ be the coordinates functions of the representative cone of V. Let h be a homogeneous element of degree m in the function field of the cone such that h/t_i^m is regular on V outside the closed set defined by $t_i=0$ for each i. Then h can be expressed as a polynomial in the t_i with coefficients in $\mathfrak{K}$ provided m is not less than a fixed integer independent of h.* Now, let $\mathfrak{L}_m$ be the linear system on C which we defined before. If we take m sufficiently large, we can find a positive divisor $\mathfrak{r}$ of C such that $\mathfrak{p}+\mathfrak{r}$ is an intersection product of C with a hypersurface H_0 of order m. We shall show that $|\mathfrak{p}|$ coincides with $\mathfrak{L}_m-\mathfrak{r}$. Let $\mathfrak{q}$ be an arbitrary member of $|\mathfrak{p}|$. Then, by definition we can find an element f of $\mathfrak{o}$ such that $\mathfrak{q}-\mathfrak{p}$ is the divisor of f in the sense of valuation theory. In other words we can find two hypersurfaces of the same order F and F_0 such that F_0 does not pass through any multiple point of C and $(F-F_0)\cdot C=\mathfrak{q}-\mathfrak{p}$ holds. If h is the function on the representative cone of C corresponding up to a constant to $F+H_0-F_0$,

then, by Zariski's lemma there exists a hypersurface H of order m such that $(F+H_0-F_0)\cdot C=\mathfrak{q}+\mathfrak{r}=H\cdot C$ holds. Therefore $\mathfrak{q}$ is a member of $\mathfrak{L}_m-\mathfrak{r}$. The converse is obvious. We note also that if $\mathfrak{p}$ is positive, we have only to take m not less than the maximum of the degree of $\mathfrak{p}$ and of the fixed integer in Zariski's lemma for $V=C$.

Now, if d is a positive integer, we shall denote by $C(d)$ the d-fold symmetric product of C. Also we shall denote by $C_0(d)$ the Chow variety of positive divisors on C of degree d. It is clear that $C_0(d)$ is an open subvariety of $C(d)$, and $C(d)$ lies in a projective space P^t of dimension $t=C_n^{d+n}-1$. The following lemma is proved, not exactly in this form, by Chow [3]:

LEMMA 1. *Every point of $C_0(d)$ is simple on $C(d)$. If W is a subvariety of $C(d)$ of dimension r, then the order of W is at least equal to* $\operatorname{ord}(C)^r$. *The extremal value* $\operatorname{ord}(C)^r$ *is attained if W is the variety of a "sufficiently general" linear system on C and also if $r=d$.*

Here a linear system $\mathfrak{L}$ is sufficiently general if the following condition is satisfied: Take a field K over which the corresponding subvariety of $C(d)$ is defined. Let $\mathfrak{p}=\sum_{i=1}^{d}P_i$ be a generic member of $\mathfrak{L}$ over K. Then all the d points $P_1,\cdots,P_d$ are generic points of the curve C over K and distinct from each other; moreover, any r of these points are independent with respect to each other over K and determine the remaining $d-r$ points uniquely.

On the other hand, it is better to state here Chow's theorem on fibre systems in a form suited for our purpose. First of all, we must recall the definition of fibre systems. Let V and W be irreducible, but possibly "incomplete varieties" in projective spaces, and let p be a function on V with values in W. We assume that p is regular on V and the point-set theoretical inverse F_y of each point y of W is an irreducible variety of the same dimension and of the same order. We then call $\{F_y\}$ a *fibre system* on V over W. If we take the set of Chow points $\{y^*\}$ of the fibres F_y, we get an irreducible, possibly incomplete variety W^* such that W is a one-to-one rational transform of W^*. This variety W^* is called the *associated variety* of the fibre system. According to a theorem of Chow [2], *a point y^* is simple on W^* if and only if the corresponding fibre F_y is simple on V.*

5. Now we shall construct the so-called generalized Jacobian variety of C as an associated variety of a certain fibre system on $C_0(d)$. The following lemma is well known if C is nonsingular:

LEMMA 2. *If $\mathfrak{a}$ is a rational divisor of C over a field K and if* $\dim|\mathfrak{a}|\geqq 0$, *then* $|\mathfrak{a}|$ *contains a rational cycle over* K. *Also if* $\mathfrak{a}$ *is of degree* d, *the subvariety of* $C(d)$ *which corresponds to* $|\mathfrak{a}|$ *is defined over* K.

Proof. By assumption there exists at least one positive 0-cycle $\mathfrak{b}$ in P^n satisfying $\mathfrak{a}\rightarrow\mathfrak{b}$. Let f be the element of $\mathfrak{o}$ such that $\mathfrak{b}-\mathfrak{a}$ is the divisor of f in the sense of valuation theory. If (f) is the usual divisor of f, we have $(f)+\mathfrak{a}\succ 0$. Since $\mathfrak{a}$ is rational over K, by a theorem of Weil [16, p. 239] f can be expressed as $f=\sum_{\alpha} c_\alpha f_\alpha$, where the c_α are linearly independent elements of $\mathfrak{K}$ over K and the f_α are functions on C defined over K satisfying $(f_\alpha)+\mathfrak{a}\succ 0$. Moreover, since the singular locus of C is a closed set over K, by the same theorem the f_α are elements of $\mathfrak{o}$. In other words, we can find an element $f'=f_\alpha$ of $\mathfrak{o}$ which is defined over K and which satisfies $(f')+\mathfrak{a}\succ 0$. If $\mathfrak{b}'$ is the sum of $\mathfrak{a}$ and the divisor of f' in the sense of valuation theory, $\mathfrak{b}'$ is a member of $|\mathfrak{a}|$. Moreover $\mathfrak{b}'$ is rational over K, and this proves the first part. The above reasoning shows also that we can find a set of linearly independent functions f_α all defined over K and forming a base over $\mathfrak{K}$ of the vector space over $\mathfrak{K}$ of elements f in $\mathfrak{o}$ satisfying $(f)+\mathfrak{a}\succ 0$. The number of f_α is then necessarily finite. Let (x) be a set of independent variables over K, and let $\mathfrak{x}$ be the sum of $\mathfrak{a}$ and the divisor of $\sum_{\alpha} x_\alpha f_\alpha$ in the sense of valuation theory. Then $|\mathfrak{a}|$ coincides with the totality of specializations of $\mathfrak{x}$ over K. However since $\mathfrak{x}$, hence also its Chow point $(\mathfrak{x})$ is rational over the regular extension $K(x)$ of K, we see that $(\mathfrak{x})$ has a locus over K. It is clear that this locus is the subvariety of $C(d)$ which corresponds to $|\mathfrak{a}|$. This proves the second part.

Now, let g be the *arithmetic genus* of C. We fix an integer d by $d>2g-2$, and we shall imitate the method of Chow [3] to construct the generalized Jacobian variety of C. For that purpose we shall first verify the following assertion: Let $\mathfrak{m}$ be a positive divisor of C of degree d the components of which are independent generic points of C over K_0. Then the complete linear system $|\mathfrak{m}|$ is sufficiently general in the sense stated next to Lemma 1. In fact, if $\mathfrak{M}$ is the subvariety of $C(d)$ which corresponds to $|\mathfrak{m}|$, then, by Lemma 2, the Chow points z of $\mathfrak{M}$ is rational over $K_0(\mathfrak{m})$. In other words $K_0(z)$ is contained in $K_0(\mathfrak{m})$. However, since $K_0(\mathfrak{m})$ is of maximal dimension d over K_0, we conclude that $(\mathfrak{m})$ is a generic point of $\mathfrak{M}$ over $K_0(z)$. In particular, the linear system $|\mathfrak{m}|$ is sufficiently general. As a consequence of this, and of Lemma 1, the order of $\mathfrak{M}$ is equal to $\mathrm{ord}(C)^r$. Moreover every specialization of $\mathfrak{M}$ over K_0 is a subvariety of $C(d)$.

On the other hand, we must also verify that any specialization of a complete linear system is again a complete linear system. We remark here that our complete linear systems are defined by stricter equivalence relation than the usual one. We take an integer m not less than d and not less than the fixed integer in Zariski's lemma such that $\dim \mathfrak{L}_m = \mathrm{ord}(C) \cdot m - g$. According to the Riemann-Roch theorem for C [14], if $\mathfrak{p}$ is a divisor of C of degree d, we have $\dim |\mathfrak{p}| = d - g$. Assume that $\mathfrak{p}$ is positive, and let $\mathfrak{P}$ be the subvariety of $C(d)$ which corresponds to $|\mathfrak{p}|$. Then $\mathfrak{P}$ is the unique specialization of $\mathfrak{M}$ over the specialization $\mathfrak{m} \to \mathfrak{p}$ with reference to K_0. This is the precise formulation of what we have stated above. Let $\mathfrak{M}'$ be a specialization of $\mathfrak{M}$ over the specialization $\mathfrak{m} \to \mathfrak{p}$ with reference to K_0. If we project each point of $\mathfrak{m}$ and $\mathfrak{p}$ from the same generic subspace of dimension $n-2$ in P^n over $K = K_0(\mathfrak{M}, \mathfrak{M}', \mathfrak{m}, \mathfrak{p})$, thus obtaining two hypersurfaces of order d passing through $\mathfrak{m}$ and $\mathfrak{p}$, and if we add to them another generic hypersurface of order $m-d$ which is independent over K to the above subspace of P^n, we get two hypersurfaces H and H' of order m in P^n. It follows from the construction that H' is a specialization of H over the specialization $(\mathfrak{M}, \mathfrak{m}) \to (\mathfrak{M}', \mathfrak{p})$ with reference to K_0. We put $\mathfrak{r} = C \cdot H - \mathfrak{m}$ and $\mathfrak{r}' = C \cdot H' - \mathfrak{p}$. It is then clear that $\mathfrak{r}$ and $\mathfrak{r}'$ are positive divisors on C such that $|\mathfrak{m}| = \mathfrak{L}_m - \mathfrak{r}$ and $|\mathfrak{p}| = \mathfrak{L}_m - \mathfrak{r}'$. Since $\mathfrak{M}'$ is a specialization of $\mathfrak{M}$ over the specialization $\mathfrak{r} \to \mathfrak{r}'$ with reference to K_0, it is carried by $\mathfrak{P}$. However, since $\mathfrak{M}'$ is irreducible and of the same dimension as $\mathfrak{P}$, we get $\mathfrak{M}' = \mathfrak{P}$.

From now on, Chow's argument [3] can be taken over verbatim to the present case. Hence we are satisfied with outlining his method by referring to the original paper of Chow for details of the proof: Let J be the locus of the Chow point z of $\mathfrak{M}$ over K_0. Then the set of Chow points of the varieties $\mathfrak{P}$ of complete linear systems on C of degree d is a subset J_0 of J. Moreover, if $(\mathfrak{p})$ is a point of $C_0(d)$ and if $\Psi(\mathfrak{p})$ denotes the Chow point of the variety $\mathfrak{P}$ which corresponds to $|\mathfrak{p}|$, then Ψ is a function on $C_0(d)$ having K_0 as a field of definition. This follows from Lemma 2. Also Ψ is "continuous" in the sense that it is commutative with specializations. If we remember the nonsingular character of $C_0(d)$, we can conclude that Ψ is regular on $C_0(d)$.

Finally, we shall show that J_0 is an open subvariety of J over K_0. Since the singular locus of C is a closed set over K_0, we conclude that $C(d) - C_0(d)$ is a closed set over K_0. We note also that every component of $C(d) - C_0(d)$ is biregularly equivalent to $C(d-1)$. In particular such a component is of

dimension $d-1$, hence we can attach a positive $(d-1)$-cycle X in P^t which is rational over K_0 and which has the same components as $C(d)-C_0(d)$. Now let $\mathfrak{P}$ be a subvariety of $C(d)$ which corresponds to a point of J. Then, by a theorem in the theory of associated forms [1], the condition that $\mathfrak{P}$ is contained in $C(d)-C_0(d)$ can be expressed by a set of doubly homogeneous equations in the Chow coordinates of $\mathfrak{P}$ and the Chow coordinates of X with coefficients in the prime field, and this gives a set of homogeneous equations in the Chow coordinates of $\mathfrak{P}$ with coefficients in K_0. This set of equations defines a closed subset of J over K_0, and J_0 is its complement.

In conclusion, the totality of $\mathfrak{P}\cap C_0(d)$ forms a fibre system on $C_0(d)$ with J_0 as its associated variety. Since $C_0(d)$ is nonsingular, Chow's theorem on fibre systems guarantees the *nonsingular character* of J_0.

6. Now, we can modify d under the condition $d>2g-2$ so that C carries a rational divisor $\mathfrak{a}$ of degree d over K_0. If θ is a divisor on C of degree zero, we define $\Phi(\theta)$ to be the Chow point of the variety which corresponds to $|\theta+\mathfrak{a}|$. It then follows from Lemma 2 that $\Phi(\theta)$ is rational over $K_0(\theta)$. Moreover, every point of J_0 can be written as $\Phi(\theta)$ with some θ. Since the set of divisors on C of degree zero forms a group, we can introduce an abstract group structure in J_0 so that Φ becomes a homomorphism. Then the law of composition in J_0 is "continuous" in the sense that it is commutative with specializations. Also, by Lemma 2 law is normal over K_0 in the sense of Weil [17, p. 51]. Therefore, from the nonsingular character of J_0 we can conclude that the "naive group operation" in J_0 is an actual group operation. In other words J_0 turns out to be a commutative group variety over K_0. Here we have $\Phi(\theta)=0$ if and only if $\theta\to 0$. The group of all such θ is a homomorphic image of the group of units in $\mathfrak{o}$, the kernel being the multiplicative group of $\mathfrak{K}$. Therefore Φ gives a rational homomorphism over K_0 [3] of the divisor group on C of degree zero onto J_0 such that the kernel is the group of principal divisors on C in the above sense. Thus J_0 and Φ are uniquely determined by C up to isomorphic transformations σ: $(J_0,\Phi)\to(\sigma J_0,\sigma\circ\Phi)$. Here, is is better to remark that J_0 is determined together with its projective embedding by the curve C in P^n and the integer d, while the group structure in J_0 as well as the homomorphism Φ depend also on the choice of the reference divisor $\mathfrak{a}$ of C. If we do not want to define the group structure *over* K_0, we can take any divisor of degree d as a reference divisor.

We shall show finally that J_0 is the generalized Jacobian variety of C in the sense of Rosenlicht. Let K be an extension of K_0 over which C carries a

rational divisor A of degree one. We can then define over K the canonical function ϕ of C by $\phi(M) = \Phi(M - A)$ for any simple point M of C. Let $M_1, \cdots, M_g$ be independent generic points of C over K and put $z = \sum_{i=1}^{g} \phi(M_i)$. Also, let $\mathfrak{m}^* = \sum_{i=1}^{d} M_i^*$ be a generic member of $|\sum_{i=1}^{g} M_i - g \cdot A + \mathfrak{a}|$ over $F = K(\sum_{i=1}^{g} M_i)$. Then, first of all, z is rational over F. Also, $\sum_{i=1}^{g} M_i$ determines a complete linear system of dimension zero [15]. Therefore F is a subfield of $K(\mathfrak{m}^*)$ by Lemma 2, and F is purely inseparable over $K(z)$. Since $K(\mathfrak{m}^*)$ is regular over $K(z)$, we conclude $F = K(z)$. This shows that our group variety J_0 is isomorphic to the generalized Jacobian variety of C with $\mathfrak{o}$ as its reference semi-local ring [15]. We state some of our results in the following way:

THEOREM 2. *Let C be an arbitrary irreducible curve in P^n having K_0 as the smallest field of definition. Then we can construct its generalized Jacobian variety J_0 over K_0 in a projective space P^N. The construction depends on a positive integer d, but once d is fixed, it is unique.*

In the Appendix we shall determine the "linear equivalence class" of the hyperplane sections of J_0, thus revealing the nature of the construction.

7. We shall treat more in detail the case when C has one and only one ordinary double point. Let K be the smallest algebraically closed field of definition for C and for the ambient surface. If C^* is the normalization of C over K, then C^* is nonsingular. We shall show that the genus γ of C^* is equal to $g - 1$. Let $\mathfrak{o}$ be the local ring of C at the double point, say Q. Since Q is an ordinary double point, the completion $\mathfrak{o}^*$ of $\mathfrak{o}$ consists of pairs of formal power series with the same constant terms. Therefore, if $\mathfrak{O}$ is the integral closure of $\mathfrak{o}$, its completion $\mathfrak{O}^*$ with respect to the semi-local topology is a direct sum of two formal power series rings. In particular, the factor module $\mathfrak{O}^*/\mathfrak{o}^*$ is a vector space of dimension one over $\mathfrak{K}$. Since the vector space $\mathfrak{O}/\mathfrak{o}$ is canonically isomorphic to $\mathfrak{O}^*/\mathfrak{o}^*$, the space $\mathfrak{O}/\mathfrak{o}$ is also of dimension one. However this dimension is equal to $g - \gamma$ [14], hence $g = \gamma + 1$. The above argument implies also that Q corresponds to two distinct points Q_1^* and Q_2^* of C^*. Moreover, a function f on C^* which is regular at Q_1^* and Q_2^* belongs to $\mathfrak{o}$ if and only if $f(Q_1^*) = f(Q_2^*)$ holds. This fact will be used presently.

Now, we shall denote by $C(d-1, Q)$ the difference $C(d) - C_0(d)$. Also, for $\alpha = 1, 2$, we shall denote by $C^*(d-1, Q_\alpha^*)$ the set of Chow points of

positive divisors on C^* of degree d containing Q_α^* as a component. It is clear that $C(d-1, Q)$ is biregularly equivalent to $C(d-1)$. Similarly $C^*(d-1, Q_\alpha^*)$ is biregularly equivalent to $C^*(d-1)$. Moreover, $C^*(d)$ is nonsingular and it is the normalization of $C(d)$ over K.

Let $M_1, \cdots, M_d$ be independent generic points of C over K, and put $\mathfrak{m} = \sum_{i=1}^{d} M_i$. Let $\mathfrak{m}^* = \sum_{i=1}^{d} M_i^*$ be the unique transform of $\mathfrak{m}$ on C^*. Then, the complete linear system $|\mathfrak{m}|$ is transformed to a linear subsystem of $|\mathfrak{m}^*|$. Let $\mathfrak{M}^{**}$ be the subvariety of $C^*(d)$ of dimension $d-g$ which corresponds to this linear subsystem. Also, let $\mathfrak{M}^*$ be the subvariety of $C^*(d)$ which corresponds to $|\mathfrak{m}^*|$. Then, the Chow point z^* of $\mathfrak{M}^*$ has a locus J^* over K, and J^* is the Jacobian variety of C^* [3]. We shall show that the locus J^{**} of the Chow point z^{**} of $\mathfrak{M}^{**}$ is birationally equivalent to J over K. Let $\mathfrak{M}$ be the subvariety of $C(d)$ which corresponds to $|\mathfrak{m}|$, and let z be the Chow point of $\mathfrak{M}$. Then z is purely inseparable over $K(z^{**})$ and also z^{**} is purely inseparable over $K(z)$. Since $K(\mathfrak{m}^*) = K(\mathfrak{m})$ is regular over $K(z^{**})$ and over $K(z)$, we get $K(z^{**}) = K(z)$. This proves our assertion. In particular, we can apply Lemma 1 to $\mathfrak{M}^{**}$, and $\mathfrak{M}^{**}$ is of order equal to $\operatorname{ord}(C^*)^{d-g}$. Moreover, every specialization of $\mathfrak{M}^{**}$ over K is a subvariety of $C^*(d)$.

Now, let $(\mathfrak{P}, \mathfrak{P}^{**})$ be an arbitrary specialization of $(\mathfrak{M}, \mathfrak{M}^{**})$ over K. Then, this is the unique extension of $\mathfrak{M}^{**} \to \mathfrak{P}^{**}$ over K, and also of $\mathfrak{M} \to \mathfrak{P}$ over K if $\mathfrak{P}$ is not contained in $C(d-1, Q)$. This follows from the fact that $C^*(d)$ is the normalization of $C(d)$ over K. We shall examine the case when $\mathfrak{P}$ is contained in $C(d-1, Q)$. Since linear equivalence is preserved under specializations, $\mathfrak{P}^{**}$ is contained in a variety $\mathfrak{P}^*$ of a complete linear system of degree d on C^*, say $\mathfrak{G}$. Since $d > 2\gamma$, we have $\dim \mathfrak{G} = d - \gamma = d - g + 1$. Moreover $\mathfrak{P}^* \cap C^*(d-1, Q_\alpha^*)$ are the varieties corresponding to $(\mathfrak{G} - Q_\alpha^*) + Q_\alpha^*$ for $\alpha = 1, 2$. Since $d - 1 > 2\gamma - 1$, these linear systems are both of dimension $d - g$. Since the variety $\mathfrak{P}^{**}$ is contained in the union of $C^*(d-1, Q_\alpha^*)$, it must coincide with one of the $\mathfrak{P}^* \cap C^*(d-1, Q_\alpha^*)$. We note that these two varieties are distinct, because $\mathfrak{G} - Q_\alpha^*$ have no fixed points. We shall now show that the totality of $\mathfrak{P}^{**}$ induces a fibre system on $C^*(d) - \bigcap_\alpha C^*(d-1, Q_\alpha^*)$ with J^{**} as its associated variety. We shall first show that the fibres have no common point. Let $(\mathfrak{q}^*)$ be a point of $C^*(d)$ not in $\bigcap_\alpha C^*(d-1, Q_\alpha^*)$. If $\mathfrak{q}^*$ is free from Q_1^* and Q_2^*, there is no ambiguity. Suppose that $\mathfrak{q}^*$ contains, say Q_1^*. Let $\mathfrak{P}^{**}$ be a fibre passing through $(\mathfrak{q}^*)$. Then $\mathfrak{P}^{**}$ must be contained in

$C^*(d-1, Q_1^*)$. Otherwise, since $\mathfrak{P}^{**}$ can not be contained in $C^*(d-1, Q_2^*)$, the corresponding subvariety $\mathfrak{P}$ of $C(d)$ will be a variety of a complete linear system on C. Thus by a remark we made in the beginning of this section, $\mathfrak{q}^*$ must contain Q_2^*, which is not the case. We conclude that $\mathfrak{P}^{**}$ is the variety of $(|\mathfrak{q}^*|-Q_1^*)+Q_1^*$, and this shows that in any case $(\mathfrak{q}^*)$ determines the fibre $\mathfrak{P}^{**}$ uniquely. The correspondence $(\mathfrak{q}^*)\to\mathfrak{P}^{**}$ defines a single-valued map from $C^*(d)-\bigcap_\alpha C^*(d-1, Q_\alpha^*)$ onto J^{**}. Since we know that this map is a function on $C^*(d)$ with K as a field of definition, and since $C^*(d)$ is nonsingular, it is regular on $C^*(d)-\bigcap_\alpha C^*(d-1, Q_\alpha^*)$. This completes our proof.

Since $C^*(d)$ is nonsingular, by Chow's theorem J^{**} is nonsingular. From this and from what we remarked before, it follows that the birational correspondence between J and J^{**} is regular on J^{**} and biregular on J_0. On the other hand, there is a function p on J^{**} defined over K with values in J^*. In fact, using our previous notation $K(\mathfrak{m}^*)$ is regular over $K(z^{**})$, while z^* is rational over $K(\mathfrak{m}^*)$ and purely inseparable over $K(z^{**})$. Hence p can be defined over K by $p(z^{**})=z^*$. Since p is single-valued on J^{**} and since J^{**} is nonsingular, p is regular on J^{**}. For $\alpha=1,2$, the Chow points of varieties corresponding to linear systems of the form $(\mathfrak{G}-Q_\alpha^*)+Q_\alpha^*$ form a subvariety J_α^*. Since J_α^* is the associated variety of the fibre system of the varieties of complete linear systems on $C^*(d-1, Q_\alpha^*)$, it is biregularly equivalent to J^*. Actually p induces such a biregular map on J_α^*. We note that J_1^* and J_2^* are disjoint. Now, by a similar reasoning we conclude that $J-J_0$ is biregularly equivalent to J^*. Actually, the birational correspondence between J and J^{**} induces such a biregular map on J_α^*. Thereby, if z_α^* are points of J_α^* for $\alpha=1,2$, the images of z_1^* and z_2^* on $J-J_0$ coincide if and only if $p(z_1^*)-p(z_2^*)=\phi^*(Q_1^*)-\phi^*(Q_2^*)$ holds. Here ϕ^* is the canonical function on C^*. We summarize some of our results in the following way:

SUPPLEMENT. *If C has one and only one ordinary double point, the normalization J^{**} of J is nonsingular. Moreover, if J^* is the Jacobian variety of the nonsingular model C^* of C, then $J-J_0$ is biregularly equivalent to J^* and is a "double variety" of J.*

Also, we can see easily that $p^{-1}(z^*)$ is biregularly equivalent over $K(z^*)$ to a projective straight line for every z^* on J^*. We can derive from this that J^{**} is the projective line bundle over J^* which is obtained by "completing" the affine line bundle $J_0=J^{**}-J_1^*-J_2^*$ over J^*. Moreover the

invariant of this bundle in the sense of Weil [18] is given by the class of linear equivalence of $\Theta^*_{\phi^*(Q_1^*)} - \Theta^*_{\phi^*(Q_2^*)}$. Here Θ^* is the divisor of J^* which is defined as the transform on J^* of the $(\gamma - 1)$-fold product of C^* by ϕ^*.

3. Néron varieties of a nonsingular surface.

8. Let C and C' be two irreducible curves in P^n such that C' is a specialization of C over a field K. Let J and J' be the completed generalized Jacobian varieties of C and C' respectively. If C and C' have the same arithmetic genus, say g, and if we use the same reference integer d greater than $2g - 2$ in the constructions of J and J', then J and J' have the same ambient space P^N and have the same dimension. In a separate paper we proved the following lemma [5]:

LEMMA. *If $\mathfrak{L}_m$ and $\mathfrak{L}'_m$ refer to C and C' for a sufficiently large m, then a specialization of a member of $\mathfrak{L}_m$ over the specialization $C \to C'$ with reference to K is a member of $\mathfrak{L}'_m$.*

We can now prove the following "compatibility theorem."

THEOREM 3. *If C is nonsingular, and if C' is either nonsingular or has one and only one ordinary double point, then J' is the unique specialization of J over the specialization $C \to C'$ with reference to K.*

Proof. Let J'' be a specialization of J over the specialization $C \to C'$ with reference to K. Our purpose is to show that $J'' = J'$. Thereby we can replace (J, C) by its generic specializations over K and in this way we can extend K arbitrarily. Therefore we can assume from the beginning that J'' and C' are both defined over K. Let F be the smallest field of definition of C containing K. Let Γ be the graph of the function Ψ on $C(d)$ with values in J which we introduced in 5; similarly Γ' is defined for C'. We shall show that Γ' is the unique specialization of Γ over the specialization $(J, C) \to (J'', C')$ with reference to K.

Let Γ'' be a specialization of Γ over this specialization. Let L be a linear subspace of co-dimension d in the ambient space of Γ which is generic over an algebraically closed field of definition E of Γ'' containing F. Then, any point ξ' in $L \cdot \Gamma''$ is a generic point of some component of Γ'' over E. Moreover, in this way we get a generic point of every component of Γ'' over E. Let ξ be a point in $L \cdot \Gamma$ which is specialized to ξ'. Since ξ is a generic point of Γ over F, it is of the form $(\mathfrak{m}) \times z$. Here $(\mathfrak{m})$ is a generic point of $C(d)$

over F, and z is the Chow point of the variety $\mathfrak{M}$ associated with $|\mathfrak{m}|$. On the other hand, since the specialization of $C(d)$ over the specialization $C \rightarrow C'$ with reference to K is carried by $C'(d)$, and since $C(d)$ and $C'(d)$ have the same order, $C'(d)$ is the unique specialization of $C(d)$ over that specialization. Therefore Γ'' is carried by the product $C'(d) \times P^N$. In particular, we can write ξ' in the form $(\mathfrak{m}') \times z'$, where $(\mathfrak{m}')$ is a point of $C'(d)$ and z' is the Chow point of the specialization $\mathfrak{M}'$ of $\mathfrak{M}$ over the specialization $\xi \rightarrow \xi'$ with reference to K. Since $\mathfrak{M}'$ is carried by $C'(d)$, we conclude from Lemma 1, Section 2 that $\mathfrak{M}'$ is an irreducible variety. Since z' and $(\mathfrak{m}')$ are points of J'' and $\mathfrak{M}'$ respectively, we have $\dim_E z' \leqq g$ and $\dim_{E(z')} (\mathfrak{m}') \leqq d-g$. However, since $(\mathfrak{m}') \times z'$ is of dimension d over E, the inequality signs can not hold in the above relations. We shall show that ξ' is a point of Γ'. Assume first that $\mathfrak{m}'$ is not a divisor of C'. Then, it contains the double point Q of C' at least once. Since $(\mathfrak{m}')$ is a generic point of $\mathfrak{M}'$ over $E(z')$, we conclude that $\mathfrak{M}'$ is contained in $C'(d-1, Q) = C'(d) - C'_0(d)$. If C^* is the nonsingular model of C', there are only a finite number of positive divisors $\mathfrak{m}_\alpha{}^*$ on C^* having $\mathfrak{m}'$ as a projection. Let U_α be the projections on $C'(d)$ of the varieties corresponding to the complete linear systems $|\mathfrak{m}_\alpha{}^*|$ on C^*. Then $\mathfrak{M}'$ is contained in $\bigcup_\alpha U_\alpha \cap C'(d-1, Q)$ according to the previous lemma. However, this is a closed set over $K(\mathfrak{m}')$ of dimension $d-g$. Hence $\mathfrak{M}'$ must coincide with one of its components. In particular, the Chow point z' of $\mathfrak{M}'$ must be algebraic over $K(\mathfrak{m}')$, hence a fortiori over $E(\mathfrak{m}')$. Thus $(\mathfrak{m}')$ is a generic point of $C'(d)$ over E, and this is a contradiction. Therefore $\mathfrak{m}'$ must be a divisor of C'. In this case, as in 5, we can find two positive divisors $\mathfrak{r}$ and $\mathfrak{r}'$ of C and C' respectively such that $\mathfrak{r}'$ is a specialization of $\mathfrak{r}$ over the specialization $(\mathfrak{m}, \mathfrak{M}) \rightarrow (\mathfrak{m}', \mathfrak{M}')$ with reference to K and such that $|\mathfrak{m}| = \mathfrak{L}_m - \mathfrak{r}$ and $|\mathfrak{m}'| = \mathfrak{L}'_m - \mathfrak{r}'$ for a sufficiently large m. Since we know that $\mathfrak{M}'$ is irreducible, we can conclude from the previous lemma that $\mathfrak{M}'$ is the variety associated with $|\mathfrak{m}'|$. Therefore ξ' is a point of Γ'. Since ξ' is a generic point over E of an arbitrary component of Γ'', we conclude that Γ'' is an integer multiple of Γ'. However, since $\mathrm{pr}_1(\Gamma'')$ coincides with $\mathrm{pr}_1(\Gamma') = C'(d)$ as a specialization of $\mathrm{pr}_1(\Gamma) = C(d)$, we get $\Gamma'' = \Gamma'$.

Finally, let $R = (M_1, \cdots, M_{d-g})$ be a set of $d-g$ independent generic points of C over F, and let R' be an isolated specialization of R over the specialization $(J, C) \rightarrow (J'', C')$ with reference to K. Then R' is a set of $d-g$ independent generic points of C' over K. If we define W to be the subvariety of $C(d)$ whose points correspond to divisors of C containing the

$d-g$ points of R as components, W is biregularly equivalent to $C(g)$. We define W' similarly for C' and R'. It is clear that W' is the unique specialization of W over the specialization $(R, J, C) \to (R', J'', C')$ with reference to K. On the other hand, the intersection product $\Gamma \cdot (W \times P^N)$ is defined and is a birational correspondence between W and J. We shall show that the intersection product $\Gamma' \cdot (W' \times P^N)$ is also defined and is a birational correspondence between W' and J'. Let $(\mathfrak{m}') \times z'$ be a generic point of some component of the intersection $\Gamma' \cap W' \times P^N$ over the algebraic closure F of $K(R')$. Then, by the results in 7 we get $\dim_{F(\mathfrak{m}')} z' \leqq 1$ in all cases. Here the equality holds if and only if $\mathfrak{m}'$ contains the double point at least twice. Since $(\mathfrak{m}') \times z'$ is of dimension g over F, we conclude that the components of $\mathfrak{m}'$ are R' and g independent generic points of C' over F. Our assertion follows from this. Therefore, $J' = \mathrm{pr}_2[\Gamma' \cdot (W' \times P^N)]$ is a specialization of $J = \mathrm{pr}_2[\Gamma \cdot (W \times P^N)]$ over the specialization $J \to J''$ with reference to K, whence $J' = J''$. q. e. d.

9. From now on, we fix a nonsingular algebraic surface V in P^n. Let $\{C_u\}$ be a general linear pencil on V whose parametrization is defined over an algebraically closed field k. We may assume, if necessary, that k is the algebraic closure of the smallest field of definition of V. Since V is nonsingular, it is normal and hence all C_u have the same arithmetic genus, say g [22]. Also, if A is one of the base points, k being algebraically closed, A is simple on V and is rational over k. Now, let J_u be the completed generalized Jacobian variety of C_u constructed with reference to a fixed integer d greater than $2g-2$. Then J_u is defined over $k(u)$, hence there exists a subvariety $\mathcal{J}$ of the product $D \times P^N$ with k as a field of definition such that $\mathcal{J} \cdot (u \times P^N) = u \times J_u$ for all generic points u of D over k. Then, by the compatibility theorem the same formula remains valid for all u on D. If we apply the criterion of multipilicity one to this situation [16, p. 141], we can conclude that a point $u \times z$ of $\mathcal{J}$ is simple on $\mathcal{J}$ as long as z is simple on J_u. If we define a function p on $\mathcal{J}$ over k by $p(u \times z) = u$, then p is regular on $\mathcal{J}$. Therefore $\{u \times J_u\}$ is a fibre system on $\mathcal{J}$ over D. Since D is nonsingular, it is biregularly equivalent to the associated variety of the fibre system. We call $\mathcal{J}$ the *Néron variety of V associated with $\{C_u\}$*, and we state our main theorem the first part of which is trivial:

THEOREM 4. *To every general linear pencil on a nonsingular surface we can associate its Néron variety. It is the variety of the parametrization of Jacobian varieties of the members of the pencil. The singular locus of the*

Néron variety is contained in the union of singular loci of degenerate fibres. In particular, it has only negligible singularities.

Now, we can use $d \cdot A$ as the reference divisor to introduce a group structure in $(J_u)_0$ over $k(u)$. At the same time we can normalize the canonical function ϕ_u of C_u by $\phi_u(A) = 0$. Under these agreements, we can state the following assertions: (i) *If (x, y) is a pair of points of J_u which is specialized to (x', y') over the specialization $u \to u'$ with reference to k, and if both x' and y' are points of $(J_{u'})_0$, then $x' + y'$ is the unique specialization of $x + y$ over the above specialization.* (ii) *If Γ_u and $\Gamma_{u'}$ are the graphs of the normalized canonical functions on C_u and $C_{u'}$, then $\Gamma_{u'}$ is the unique specialization of Γ_u over the specialization $u \to u'$ with reference to k.*

Let M be a generic point of V over k, and let C_u be the member of the pencil passing through M. Then u is rational over $k(M)$. Moreover, if we put $x = \phi_u(M)$, we have $k(u)(M) = k(u)(x)$, i.e., $k(M) = k(u, x)$. In other words, we can define a birational map ϕ over k from V into $\mathcal{G}$ by $\phi(M) = u \times \phi_u(M)$. We note that the image of ϕ is not contained in the singular locus of $\mathcal{G}$.

10. We shall now treat the connection of Albanese varieties and linear differential forms of the first kind attached to V and $\mathcal{G}$. Here, the Albanese variety of an arbitrary variety U is a pair of an Abelian variety A and a function f on U with values in A satisfying the two conditions: (i) The image of f generates A. (ii) If h is a function on U with values in an Abelian variety B, there exists a homomorphism α from A into B such that $h = \alpha \circ f$ + constant. These two properties determine (A, f) up to isomorphic transformations, and the existence of (A, f) is proved by Chow and Matsusaka [4, 10]. On the other hand, we summarized some of the basic properties of differential forms of the first kind already elsewhere [6]. They are mainly due to Koizumi [8]. We need also a result of Kawahara [7], which, in an apparently stronger form, can be stated as follows: *Let f be a rational map of a variety V into another variety U such that the image variety is not contained in the singular locus of U. Then the associated linear map δf maps every differential form of the first kind on U to a differential form of the first kind on V.* We shall use the terminology such as injective, surjective and bijective instead of "one-to-one into," "onto" and "one-to-one onto." The mappings we consider later are linear mappings of differential forms.

We shall first consider the symmetric product $V(d)$ of an arbitrary variety V in P^n. We must get an information about the singular locus of

$V(d)$. The following lemma is a consequence of Chow's theorem [2], but we shall give a direct proof:

LEMMA 1. *If $P_1, \cdots, P_d$ are distinct simple points of V, the Chow point of $\sum_{i=1}^{d} P_i$ is simple on $V(d)$.*

Proof. Let (a_i) be the homogeneous coordinates of P_i, and let x_i be d independent generic points of V over an algebraically closed field of definition K of V over which P_i are rational. Let b and y be the Chow points of $\sum_{i=1}^{d} P_i$ and $\sum_{i=1}^{d} x_i$ respectively. Since any linear coordinates transformation in P^n corresponds to a special type of linear coordinates transformation in the ambient space of $V(d)$, we can assume the following: We can normalize (a_i) as $a_{i0} = 1$. If (x_i) is the representative of x_i satisfying $x_{i0} = 1$, then $(x_{i1}, \cdots, x_{ir})$ form a set of local coordinates of V at any P_k. If we put $D_j(X) = \prod_{d \geqq k > i} (X_{kj} - X_{ij})$, we have $D_j(a) \neq 0$ for $j = 1, \cdots, r$.

Now, let (y) be the representative of y which corresponds to the above representation. Let $(y_{1j}, \cdots, y_{dj})$ be the elementary symmetric functions of $(x_{1j}, \cdots, x_{dj})$. Then, all the y_{ij} appear as components of (y). Let $G_{ij}(X, Y) = 0$ be the set of equations for (x_{ij}) and (y_{ij}) with coefficients in the prime field. Then, $\det(\partial G_{ij}/\partial X_{kl})$ coincides with $\prod_{j=1}^{r} D_j(X)$ up to a possible change of sign. Therefore, this determinant is different from zero at $(X_{ij}) = (a_{ij})$. Since $(x_{i1}, \cdots, x_{ir})$ form a set of local coordinates of V at any P_k, we can find $n - r$ polynomials $F_\alpha(X)$ in the defining ideal of V over K such that $\det(\partial F_\alpha/\partial X_\beta)$ for $\beta = r + 1, \cdots, n$ are different from zero at $(X) = (a_k)$ for $k = 1, \cdots, d$. Therefore, we can apply a criterion of multiplicity one [16, p. 66] to conclude the following: Let $(i_1, \cdots, i_d)$ be any permutation of $(1, \cdots, d)$. Then, the specialization $(x_1, \cdots, x_d) \to (a_{i_1}, \cdots, a_{i_d})$ is of multiplicity one over the specialization $(y_{ij}) \to (b_{ij})$ with reference to K. Therefore, by a stronger reason the specialization $(y) \to (b)$ is also of multiplicity one over the specialization $(y_{ij}) \to (b_{ij})$ with reference to K. Hence, by another criterion of multiplicity one [16, pp. 127, 139], b is simple on $V(d)$.

Let $P_2, \cdots, P_d$ be $d - 1$ distinct simple points of V, and let P_1 be an arbitrary point of V. Then, the Chow point of $\sum_{i=1}^{d} P_i$ can be considered as the value of a function Ψ on V at P_1. The previous lemma implies that the

image of Ψ is not contained in the singular locus of $V(d)$. Therefore, any rational map of $V(d)$ into an Abelian variety is regular along the image of Ψ [17, p. 27]. In particular, if (A, F) is the Albanese variety of $V(d)$, then the composite function $F \circ \Psi$ is defined.

LEMMA 2. *The Albanese varieties of V and $V(d)$ are isomorphic. More precisely, $(A, F \circ \Psi)$ is the Albanese variety of V. Also, $\delta\Psi$ is injective as a mapping of linear differential forms of the first kind.*

Proof. The first part can be proved easily, hence we can omit its proof. Let V_d be the d-fold product $V \times \cdots \times V$ of V. Let θ be a linear differential form of the first kind on $V(d)$. If p is the natural projection of V_d onto $V(d)$, then $\delta p \cdot \theta$ is a linear differential form of the first kind on V_d. Therefore, if we denote by p_i the projection of V_d to its i-th factor, we have $\delta p \cdot \theta = \sum_{i=1}^{d} \delta p_i \cdot \omega_i$ with linear differential forms of the first kind ω_i on V [8]. Since $\delta p \cdot \theta$ is invariant with respect to the interchange of factors of V_d, we have $\delta p_i \cdot \omega_i = \delta p_i \cdot \omega_j$ for any i and j. Therefore, we have $\omega_i = \omega_j$ for any i and j, whence we can drop the suffices. It is clear that $\omega = \delta\Psi \cdot \theta$ holds. Since p is separable, $\theta \neq 0$ implies $\delta p \cdot \theta \neq 0$, whence $\omega \neq 0$. In other words, $\delta\Psi$ is injective.

On the other hand, the following lemma can be proved in the same way as Lemma 2. It is in fact a corollary of Lemma 2.

LEMMA 3. *Let (J, f) be the Albanese variety of a nonsingular curve C. Then δf is bijective as a mapping of linear differential forms of the first kind.*

11. We can now treat the connetcion between our nonsingular surface V and the Néron variety $\mathcal{J}$, i.e., we can prove the following theorem:

THEOREM 5. *Let (A, f) be the Albanese variety of $\mathcal{J}$. Then $(A, f \circ \phi)$ is the Albanese variety of V. Moreover, $\delta\phi$ is injective as a mapping of linear differential forms of the first kind.*

Proof. Let u be a generic point of D over k, and let $M_1, \cdots, M_g$ be independent generic points of C_u over $k(u)$. Put $\mathfrak{m} = \sum_{i=1}^{g} M_i$ and $x = \sum_{i=1}^{g} \phi_u(M_i)$. Then we have $k(\mathfrak{m}) = k(u, x)$. Therefore, a function ψ on $\mathcal{J}$ with values in $V(g)$ is defined over k such that the Chow point $(\mathfrak{m})$ of $\mathfrak{m}$ is the value of ψ at $u \times x$. We note that the image of ψ is not contained in the singular locus of $V(g)$. Now, in the statement of the theorem $f \circ \phi$ is defined since f

is regular along the image of ϕ. Let (B, h) be the Albanese variety of V. Then, there exists a homomorphism β from B into A such that $f\circ\phi=\beta\circ h$ $+$ constant. On the other hand, let $M_1,\cdots,M_g$ be distinct simple points of V. Then, we can define a function H on $V(g)$ with values in B such that $\sum_{i=1}^{g} h(M_i)$ is the value of H at the Chow point of $\sum_{i=1}^{g} M_i$. The function H is regular along the image of ψ, hence $H\circ\psi$ is defined. We can then find a homomorphism α from A into B such that $H\circ\psi=\alpha\circ f+$ constant. After these preparations, take an extension K of k over which A, B, f and h are all defined. As before, let u be a generic point of D over K, and let $M_1,\cdots,M_g$ be independent generic points of C_u over $K(u)$. Then we have

$$H\circ\psi(u\times\sum_{i=1}^{g}\phi_u(M_i))=\sum_{i=1}^{g}h(M_i), \text{ and also}$$

$$H\circ\psi(u\times\sum_{i=1}^{g}\phi_u(M_i))=\alpha\circ f(u\times\sum_{i=1}^{g}\phi_u(M_i))+\text{constant.}$$

However, since $\alpha\circ f$ induces a homomorphism of $u\times J_u$ into B up to a possible translation in B [17, p. 34], we have

$$\alpha\circ f(u\times\sum_{i=1}^{g}\phi_u(M_i))=\sum_{i=1}^{g}\alpha\circ f(u\times\phi_u(M_i))+\text{constant.}$$

Therefore, we get $\sum_{i=1}^{g}h(M_i)=\sum_{i=1}^{g}\alpha\circ f\circ\phi(M_i)+$ constant. Since h and $\alpha\circ f\circ\phi$ are both regular at the base points of $\{C_u\}$, by specializing $M_2,\cdots,M_g$ to one of these base points we get $h=\alpha\circ f\circ\phi+$ constant. Therefore,

$$\alpha\circ\beta\circ h=\alpha\circ f\circ\phi+\text{constant}=h+\text{constant,}$$

hence $\alpha\circ\beta=1$. In the same way, if we put $x=\sum_{i=1}^{g}\phi_u(M_i)$, we have

$$\beta\circ H\circ\psi(u\times x)=\beta(\sum_{i=1}^{g}h(M_i))=\sum_{i=1}^{g}\beta\circ h(M_i)=\sum_{i=1}^{g}f\circ\phi(M_i)+\text{constant}$$
$$=f(u\times\sum_{i=1}^{g}\phi_u(M_i))+\text{constant}=f(u\times x)+\text{constant,}$$

i.e., we have $\beta\circ H\circ\psi=f+$ constant. Therefore,

$$\beta\circ\alpha\circ f=\beta\circ H\circ\psi+\text{constant}=f+\text{constant,}$$

hence $\beta\circ\alpha=1$. This completes the proof of the first part.

Next we shall show that $\delta\phi$ is injective. Let ϖ be a linear differential

form of the first kind on $\mathcal{G}$ such that $\delta\phi\cdot\varpi=0$. Let K be an extension of k over which ϖ is defined, and let u be a generic point of D over K. Assume, for a moment, that $\mathrm{Tr}_{u\times J_u}\varpi\neq 0$. Then, by Lemma 3 we have

$$0=\mathrm{Tr}_{C_u}(\delta\phi\cdot\varpi)=\delta(\mathrm{Tr}_{C_u}\phi)\cdot\varpi\neq 0,$$

and this is a contradiction. Therefore, we have $\mathrm{Tr}_{u\times J_u}\varpi=0$. We may assume that u is the value of a numerical function on $\mathcal{G}$ with k as a field of definition. Let x be a generic point of J_u over $K(u)$, and let $(v_1,\cdots,v_g)$ be a set of independent variables in $K(u,x)$ over $K(u)$ such that $K(u,x)$ is separably algebraic over $K(u,v)$. Since ϖ is defined over K, we can write $\varpi(u\times x)$ in the form $fdu+\sum_{i=1}^{g}h_idv_i$ with f and h_i in $K(u,x)$. Since $\mathrm{Tr}_{u\times J_u}\varpi=0$, we have $fDu+\sum_{i=1}^{g}h_iDv_i=0$ for any derivation D in $K(u,x)$ over $K(u)$. Therefore, h_i are all zero and we get $\varpi(u\times x)=fdu$. However, u or $1/u$ can be included in a set of local coordinates on $\mathcal{G}$ everywhere outside the singular loci of degenerate fibres. Therefore, if $\varpi\neq 0$, we get $(\varpi)=(f)-2\cdot J_\infty$. Since ϖ is of the first kind, we have $(\varpi)\succ 0$. Thus, a strictly negative divisor is linearly equivalent to a positive divisor on the variety $\mathcal{G}$ with negligible singularities! But this is a contradiction. Therefore $\varpi=0$, hence $\delta\phi$ is injective.

Appendix.

In this Appendix we shall analyze Chow's construction of Jacobian varieties. The restriction to Jacobian varieties is merely a matter of simplicity. Let C be a nonsingular curve of genus g in P^n with K_0 as the smallest field of definition. Let J be the Jacobian variety of C which is constructed in P^N by Chow's method with respect to an integer d greater than $2g-2$. Our purpose is to determine the linear equivalence class of hyperplane sections of J. As before, we introduce a group structure in J with reference to a divisor $\mathfrak{a}$ on C of degree d. Let K be an extension of K_0 over which $\mathfrak{a}$ is rational and the canonical function ϕ of C is defined. If $P_1,\cdots,P_{g-1}$ are $g-1$ points of C, the set of points of J of the form $\sum_{i=1}^{g-1}\phi(P_i)$ is a subvariety Θ of J of dimension $g-1$ with K as a field of definition. If z is a point of J, then Θ_z is the image of Θ under the translation $x\rightarrow x+z$ of J. On the other hand if $\sum_i x_i$ is a hyperplane section of C, then, by Abel's theorem [17, p. 39] $c=\phi(\sum_i x_i)=\sum_i\phi(x_i)$ is independent of the choice of the hyperplane.

In particular, c is rational over K. We shall now prove the following assertion:

THEOREM. *A hyperplane section of J is linearly equivalent to $\Theta_z - \Theta + \mathrm{ord}(C)^{d-g+1}\cdot\Theta$ with*

$$z = (d-g+1)\,\mathrm{ord}(C)^{d-g}\,c - \mathrm{ord}(C)^{d-g+1}\phi(\mathfrak{a}).$$

Proof. Let (X), (Y) and (Z) be the sets of letters to describe equations in P^n, P^t and P^N respectively. Let (u) be a set of $n+1$ quantities. In general, if w is a point of a projective space, its homogeneous coordinates will be denoted by (w). Let $x^1,\cdots,x^d$ be d points in P^n, and let y be the Chow point of $\sum_{i=1}^{d} x^i$. Then, we have a relation of the form

$$\prod_{i=1}^{d}\Big(\sum_{j=0}^{n} u_j x^i{}_j\Big) = \sum_{j=0}^{t} \omega_j(u)\, y_j.$$

Here $\omega_j(u)$ are monomials of degree d in $u_0,\cdots,u_n$. Now, let L_v be a subspace of P^t of dimension $t-r$ which is defined by $\sum_{j=0}^{t} v^i{}_j Y_j = 0$ for $i=1,\cdots,r$. Let W^r be a positive cycle in P^t of order s such that the intersection product $W\cdot L_v$ is defined. Let $\sum_{\beta=1}^{s} y^\beta$ be this 0-cycle in P^t. Also, let (v^0) be a set of $t+1$ quantities. Then, if z is the Chow point of W, we have the following relation

$$\prod_{\beta=1}^{s}\Big(\sum_{j=0}^{t} v^0{}_j y^\beta{}_j\Big) = \sum_{j=0}^{N} \Omega_j(v^0,\cdots,v^r)\, z_j.$$

Here $\Omega_j(v^0,\cdots,v^r)$ are monomials of degree s in each $v^\alpha{}_0,\cdots,v^\alpha{}_t$ for $\alpha=0,\cdots,r$. After these preliminary remarks, we put $r=d-g$ and $s=\mathrm{ord}(C)^{d-g}$. Let $\sum_{j=0}^{n} u^\alpha{}_j X_j = 0$ be $r+1$ independent generic linear equations over K for $\alpha=0,\cdots,r$. They define $r+1$ hyperplanes in P^n intersecting C at $\sum_{i=1}^{d} x^{\alpha i}$. Let H be the hyperplane in P^N with the equation

$$\sum_{j=0}^{N} \Omega_j(\omega(u^0),\cdots,\omega(u^r))\, Z_j = 0.$$

We shall determine the intersection product $J\cdot H$ of J and H. We denote by $L_{\omega(u)}$ the subspace of P^t of dimension $t-r$ which is defined by $\sum_{j=0}^{t} \omega_j(u^\alpha) Y_j = 0$ for $\alpha=1,\cdots,r$. If $J\cdot H$ is not defined, J is contained

in H. We shall derive a contradiction from this assumption. Let z be a generic point of J over $K(u^0,\cdots,u^r)$, and let $\mathfrak{P}$ be the corresponding subvariety of $C(d)$. Then, the intersection product $\mathfrak{P}\cdot L_{\omega(u)}$ is defined. Let it be $\sum_{\beta=1}^{s} y^\beta$. Since z is a point of H, we have $\prod_{\beta=1}^{s}(\sum_{j=0}^{t}\omega_j(u^0)y^\beta{}_j)=0$, hence $\sum_{j=0}^{t}\omega_j(u^0)y_j=0$ for some $y=y^\beta$. However, since y is a point of $L_{\omega(u)}$, we also have $\sum_{j=0}^{t}\omega_j(u^\alpha)y_j=0$ for $\alpha=1,\cdots,r$. If $\sum_{j=1}^{d}x^i$ is the 0-cycle in P^n corresponding to y, we may assume that $\sum_{j=0}^{n}u^\alpha{}_jx^\alpha{}_j=0$ for $\alpha=1,\cdots,r$ and also $\sum_{j=0}^{n}u^0{}_jx^{r+1}{}_j=0$. In particular, $x^1,\cdots,x^{r+1}$ are independent generic points of C over $K(z)$, while $K(x^1,\cdots,x^d)$ is of dimension r over $K(z)$. This is a contradiction. Therefore $J\cdot H$ is defined. Now, let z be a generic point of some component W of $J\cdot H$ over the algebraic closure F of $K(u^0,\cdots,u^r)$. Then z is a generic point of J over $K_1=K(u^1,\cdots,u^r)$. Otherwise, (u^0) would be a set of independent variables over $K_1(z)$, which is not the case. Therefore, if $\mathfrak{P}$ is the subvariety of $C(d)$ which corresponds to z, then $\mathfrak{P}\cdot L_{\omega(u)}$ is again defined. Let it be $\sum_{\beta=1}^{s}y^\beta$. Then, some $y=y^\beta$ corresponds to a 0-cycle $\sum_{i=0}^{d}x^i$ such that $\sum_{j=0}^{n}u^\alpha{}_jx^\alpha{}_j=0$ for $\alpha=1,\cdots,r$ and $\sum_{j=0}^{n}u^0{}_jx^{r+1}{}_j=0$. We note that $x^1,\cdots,x^{r+1}$ are rational over F, and z is rational over $F(y)$. Therefore $x^{r+2},\cdots,x^d$ are independent generic points of C over F. Since W is the locus of the point z over F, it is of the form $\Theta_{z_\gamma-\phi(a)}$ with $z_\gamma=\sum_{i=1}^{r+1}\phi(x^i)$. Conversely, if z_γ is as above, $\Theta_{z_\gamma-\phi(a)}$ is obviously a component of $J\cdot H$. We shall now show that H is transversal to J along $\Theta_{z_\gamma-\phi(a)}$, i.e., at z. We know that $\sum_{j=0}^{N}\Omega_j(\omega(u^0),\cdots,\omega(u^r))z_j=0$ is the only relation between $(u^0),\cdots,(u^r)$ with coefficients in $K(z)$. Also, if we put $r+1$ sets of indeterminates $(U^0),\cdots,(U^r)$ instead of $(u^0),\cdots,(u^r)$, then $\sum_{j=0}^{N}\Omega_j(\omega(U^0),\cdots,\omega(U^r))z_j$ is irreducible over $K(z)$. In fact, we have

$$\sum_{j=0}^{N}\Omega_j(\omega(U^0),\omega(u^1),\cdots,\omega(u^r))z_j=\prod_{\beta=1}^{s}\prod_{i=1}^{d}(\sum_{j=1}^{n}U^0{}_jx^{\beta i}{}_j).$$

Here $(x^{\beta 1},\cdots,x^{\beta d})$ for a complete set of conjugates over $K_1(y^\beta)=K_1(z)(y^\beta)$ for each β, and $(y^1,\cdots,y^s)$ form a complete set of conjugates over $K_1(z)$, whence our assertion. Suppose now that H is not transversal to J at z. Let (z) be normalized by $z_0=1$. Then, we can find $N-g$ polynomials

$F_\lambda(Z)$ in the defining ideal of J over K_0 such that the matrices $(\partial F_\lambda/\partial z_j)$ and

$$\begin{pmatrix} \partial F_\lambda/\partial z_j \\ \Omega_j(\omega(u^0),\cdots,\omega(u^r)) \end{pmatrix}$$

have the same rank $N-g$. We thus get a relation of the form

$$\sum_{j>0} A_j(z)\Omega_j(\omega(u^0),\cdots,\omega(u^r))=0$$

with $A_j(z)$ in $K_0(z)$. However since $\sum_{j=0}^{N} z_j\Omega_j(\omega(U^0),\cdots,\omega(U^r))$ is irreducible over $K(z)$, this must divide $\sum_{j>0} A_j(z)\Omega_j(\omega(U^0),\cdots,\omega(U^r))$. On the other hand, by taking a suitable coordinates transformation in P^n if necessary, we may assume that $\Omega_0(\omega(U^0),\cdots,\omega(U^r))$ is equal to $(U^0{}_0\cdots U^r{}_r)^{ds}$. We are thus led to a contradiction. We have thus shown that $J\cdot H$ is of the form $\sum_\gamma \Theta_{z_\gamma-\phi(a)}$, where the summation is extended over distinct $\Theta_{z_\gamma-\phi(a)}$. However, since the correspondence $z\to\Theta_z$ is one-to-one [17, p. 76], the summation is extended over the distinct z_γ. Here, two distinct sets $(x^1,\cdots,x^{r+1})$ and $(x'^1,\cdots,x'^{r+1})$ correspond to different z. Otherwise we get $g=0$, and in this case everything is trivial. Thus, the number of γ is equal to $\operatorname{ord}(C)^{r+1}$, whence

$$\sum_\gamma \Theta_{z_\gamma-\phi(a)}\sim\Theta_{\sum_\gamma z_\gamma-\sum_\gamma\phi(a)}+(\operatorname{ord}(C)^{r+1}-1)\cdot\Theta$$

holds [17, p. 106]. It is now a simple matter to verify that

$$\sum_\gamma z_\gamma-\sum_\gamma\phi(a)=(d-g+1)\operatorname{ord}(C)^{d-g}c-\operatorname{ord}(C)^{d-g+1}\phi(a).$$

HARVARD UNIVERSITY AND KYOTO UNIVERSITY, JAPAN.

REFERENCES.

[1] W. L. Chow and B. L. van der Waerden, "Ueber zugeordnete Formen und algebraische Systeme von algebraische Mannigfaltigkeiten," *Mathematische Annalen*, vol. 113 (1937), pp. 692-704.

[2] W. L. Chow, "Algebraic systems of positive cycles in an algebraic variety," this JOURNAL, vol. 72 (1950), pp. 247-283.

[3] ———, "The Jacobian variety of an algebraic curve," this JOURNAL, vol. 76 (1954), pp. 453-476.

[4] ———, "Abelian varieties over function fields," *Transactions of the American Mathematical Society*, vol. 78 (1955), pp. 253-275.

[5] J. Igusa, "Arithmetic genera of normal varieties in an algebraic family," *Proceedings of the National Academy of Sciences*, vol. 41 (1955), pp. 34-37.

[6] ———, "A fundamental inequality in the theory of Picard varieties," *ibid.*, pp. 317-320.

[7] Y. Kawahara, "On the differential forms on algebraic varieties," *Nagoya Mathematical Journal*, vol. 4 (1952), pp. 73-78.

[8] S. Koizumi, "On the differential forms of the first kind on algebraic varieties," I, *Journal of the Mathematical Society of Japan*, vol. 1 (1949), pp. 273-280; II, *ibid.*, vol. 2 (1951), pp. 267-269.

[9] T. Matsusaka, "The theorem of Bertini on linear systems in modular fields," *Kyoto Mathematical Memoirs*, vol. 26 (1950), pp. 51-62.

[10] ———, "On the algebraic construction of the Picard variety," II, *Japanese Journal of Mathematics*, vol. 22 (1952), pp. 51-62.

[11] Y. Nakai, "Note on the intersection of an algebraic variety with the generic hyperplane," *Kyoto Mathematical Memoirs*, vol. 26 (1951), pp. 185-187.

[12] A. Néron, "Problèmes arithmétiques et géométriques rattachés à la notion du rang d'une courbe algébrique dans un corps," *Bulletin de la Société Mathématique de France*, vol. 80 (1952), pp. 101-166.

[13] A. Néron and P. Samuel, "La variété de Picard d'une variété normales," *Annales de l'Institut Fourier*, vol. 4 (1953), pp. 1-30.

[14] M. Rosenlicht, "Equivalence relations on algebraic curves," *Annals of Mathematics*, vol. 56 (1952), pp. 169-191.

[15] ———, "Generalized Jacobian varieties," *ibid.*, 59 (1954), pp. 505-530.

[16] A. Weil, *Foundations of algebraic geometry*, New York (1946).

[17] ———, *Variétés Abéliennes et courbes algébriques*, Paris (1948).

[18] ———, "Fibre spaces in algebraic geometry," mimeographed notes of Chicago conference on algebraic geometry and algebraic number theory (1949).

[19] O. Zariski, *Algebraic surfaces*, New York (1948).

[20] ———, "Some results in the arithmetic theory of algebraic varieties," this Journal, vol. 61 (1939), pp. 249-294.

[21] ———, "Pencils on an algebraic variety and a new proof of a theorem of Bertini," *Transactions of the American Mathematical Society*, vol. 50 (1941), pp. 48-70.

[22] ———, "Complete linear systems on normal varieties and a generalization of a lemma of Enriques-Severi," *Annals of Mathematics*, vol. 55 (1952), pp. 552-592.

FULLY REDUCIBLE SUBGROUPS OF ALGEBRAIC GROUPS.*

By G. D. Mostow.[1]

Section 1. Introduction.

In the theory of Lie groups one often encounters statements which can be asserted for *connected* groups on the strength of the appropriate analogue for Lie algebras, but which may be invalid for non-connected groups. A good example of this phenomenon is the theorem of Sophus Lie that a solvable linear Lie algebra (over a field of characteristic 0) and hence a *connected* solvable linear Lie group can be simultaneously triangularized (upon extending the ground of its algebraic closure). This powerful principle is not valid for general solvable linear groups. In fact, an examination of proofs of results that are valid for *only* connected Lie groups often reveals that Lie's theorem has been used in an essential way.

On the other hand, there are some results which are deduced for connected groups from their Lie algebras which continue to ring true when the hypothesis of connectedness is dropped. In this paper we obtain several results of such a type. The results are related for most part to properties of fully reducible groups of linear transformations. Our central result concerns a decomposition of algebraic groups which is closely related to the Wedderburn decomposition of an associative algebra into the semi-direct sum of a semi-simple subalgebra and the radical.

Our decomposition for algebraic groups can be viewed as a result on group extensions. Any algebraic group is a finite extension of the connected component of its identity G_0. To what extent is the extension splittable? Put in another way, how large a normal *connected* subgroup N can we find so that N admits a complementary subgroup? The answer in any particular case depends on the arithmetic properties of the ground field. Nevertheless for a general ground field of characteristic zero, something can still be asserted.

THEOREM. *Let N be the set of unipotent elements (i. e. eigenvalues are*

* Received June 25, 1954; revised August 4, 1955.

[1] Performed in part under Contract DA-36-034-ORD 1374 with Officer of Ordnance Research.

all 1) *of the radical of the algebraic group* G. *Let* M *be any maximal fully reducible subgroup of* G. *Then* N *is a connected normal subgroup and*

$$G = M \cdot N \, (\textit{semi-direct}).$$

For a general ground field this choice of N is the best possible.

THEOREM. *Any two maximal fully reducible subgroups of an algebraic group are conjugate under an inner automorphism.*

There is of course a similar result for algebraic Lie algebras. As an application we prove the following

THEOREM. *Let* G *be an algebraic group and let* $\mathfrak{E}$ *denote its enveloping associative algebra. Then there is a Wedderburn decomposition* $\mathfrak{S} + \mathfrak{T}$ *for* $\mathfrak{E}$ ($\mathfrak{S}$ *semi-simple,* $\mathfrak{T}$ *the radical*) *such that*

a) $\mathfrak{S} \cap G$ *is a maximal fully reducible subgroup of* G;

b) $(I + \mathfrak{T}) \cap G$ *is the subgroup of unipotent elements in the radical of* G, I *being the identity*;

c) $\mathfrak{T} \cap \mathfrak{G}$ *is the ideal of nilpotent elements in the radical of* $\mathfrak{G}$, *the Lie algebra of* G.

There is an analogous result for algebraic Lie algebras. We prove also

THEOREM. *A fully reducible group of automorphisms of a Lie algebra keeps a maximal semi-simple subalgebra invariant.*

THEOREM. *A fully reducible group of automorphisms of a solvable Lie algebra keeps a Cartan subalgebra invariant.*

In addition, there are analogous theorems for automorphisms of Lie groups. Some of the intermediate results have independent interest. For example (cf. Proposition 3.2), any rational representation of an algebraic group carries fully reducible subgroups into fully reducible subgroups. As a consequence, it can be asserted that the First Main Theorem of the Theory of Invariants is valid for fully reducible groups.

Another consequence of this same principle is the following

THEOREM. *A fully reducible group of automorphisms of a Lie or associative or Jordan algebra keeps a maximal semi-simple subalgebra invariant* (Section 5).

In conclusion, it may be noted that fully reducible subgroups behave

very much like compact subgroups of Lie groups. Indeed there is a very close relation between these two notions and we describe the relation elsewhere.

Section 2. Preliminaries.

2.1. Terminology. We deal throughout with a ground field K of characteristic 0. We employ the Zariski topology for subsets of finite dimensional linear spaces, a closed set being by definition an algebraic manifold. The terms closure, relatively open, and connected are defined in terms of closed sets in the usual way. By an endomorphism of a linear space V is meant a linear map of V into itself. By a linear group we mean a group of automorphisms of a finite dimensional linear space. The linear space of endomorphisms of a linear space V is denoted by $E(V)$.

An *algebraic group* is a group of automorphisms of a linear space V which is relatively open in its closure in $E(V)$; or, what is the same, it is a group constituting the totality of automorphisms in some algebraic manifold in $E(V)$ (cf. [2]). By the *algebraic group hull* of a set S of automorphisms is meant the intersection of all algebraic groups containing S. It is the same as the closure of S in the set of automorphisms of V, if S is a group. The connected component of the identity G_0 of an algebraic group has as its associated ideal (of polynomial functions on $E(V)$ which vanish on G_0) a prime ideal and is a normal relatively open subgroup of finite index. It should be noted that if K is the field of real numbers, an algebraic group connected in the Zariski topology need not be connected in the euclidean topology; however for K the field of complex numbers, the two notions agree.

An endomorphism T of a linear space V defines by right translation a left invariant infinitesimal transformation T^* on $E(V)$; via the derivation $P(X) \rightarrow dP(X+sXT)/ds$ at $s=0$ of the ring of polynomial functions $P(X)$ defined on $E(V)$. The *Lie algebra* of an algebraic group G is the totality of elements T in $E(V)$ such that T^* keeps invariant the ideal of polynomial functions vanishing on G (cf. [2]). An *algebraic Lie algebra* is by definition the Lie algebra of an algebraic group. An algebraic group and its connected component of the identity have the same Lie algebra. Algebraic Lie algebras are related to connected algebraic groups in much the same way as in the classical case real Lie algebras are related to connected Lie groups. In particular, a connected algebraic group keeps a linear subspace invariant if and only if its Lie algebra does; it is therefore fully reducible if and only if its Lie algebra is.

If T is an endomorphism of a linear space V, we denote by $V_b(T)$ the

totality of elements of V annihilated by some power of $T - bI$ where b is in K and I denotes the identity. We call $V_0(T)$ the nilspace of T and call T nilpotent if $V_0(T) = V$. If F is a family of endomorphisms of V, we mean by $V_0(F)$, the nilspace of F, the intersection of the nilspaces of the elements of F. For any endomorphism T there exist a unique fully reducible endomorphism s and a unique nilpotent endomorphism n such that $sn = ns$ and $s + n = T$. This decomposition is called "the Jordan sum decomposition of T," s and n being called the fully reducible and nilpotent parts of T respectively.

If T is an automorphism of the linear space V, then T has a unique product decomposition $T = s \cdot u$ where $s \cdot u = u \cdot s$, s is the fully reducible part of T and u is *unipotent*, i.e. $V_1(u) = V$. Any algebraic group contains the Jordan product components of each of its elements, and correspondingly every algebraic Lie algebraic contains the Jordan sum components of its elements. Algebraic Lie algebras are thus splittable in the sense of Malcev (cf. [7]).

Let F be a family of endomorphisms of a linear space V, and let $G(V)$ denote the group of automorphisms of V. By a *similarity* of F is meant the restriction to F of a mapping $f \to xfx^{-1}$ where x is in $G(V)$ and $xFx^{-1} = F$. A group G of similarities of F is called pre-fully reducible if there exists a fully reducible subgroup G_1 in $G(V)$ with G the restrictions to F of the similarities from G_1.

Let F be a family of endomorphisms of a linear space V which keeps a linear subspace W invariant. We denote by F_W (resp. $F_{V/W}$) the induced family of transformations of W (resp. V/W) and call these the W-part and V/W part of F respectively. If ρ is a representation of a Lie algebra $\mathfrak{G}$, we call $\mathfrak{G}$ ρ-reductive if $\rho(\mathfrak{G})$ is a fully reducible set. We denote by ad the adjoint representation of a Lie algebra, $\operatorname{ad} x(y)$ being by definition $[x, y]$.

2. 2. Cartan subalgebras. Let $\mathfrak{G}$ be a Lie algebra, let $\mathfrak{G}^0 = \mathfrak{G}$, let $\mathfrak{G}^n = [\mathfrak{G}^{n-1}, \mathfrak{G}]$, and let $\mathfrak{G}^\infty = \bigcap_n \mathfrak{G}^n$. $\mathfrak{G}$ is called nilpotent if $\mathfrak{G}^\infty = (0)$. By Engel's Theorem, $\mathfrak{G}$ is nilpotent if $\operatorname{ad} x$ is nilpotent for all x in $\mathfrak{G}$. A Cartan subalgebra of a Lie algebra $\mathfrak{G}$ is any nilpotent subalgebra $\mathfrak{H}$ such that $\mathfrak{H}$ is the nilspace of (the $\mathfrak{G}$ part of) $\operatorname{ad} \mathfrak{H}$. Any representation of $\mathfrak{G}$ which vanishes on a Cartan subalgebra vanishes on $\mathfrak{G}$. Any maximal abelian ad reductive subalgebra of a semi-simple Lie algebra $\mathfrak{G}$ is a Cartan subalgebra and vice versa. If $\mathfrak{H}$ is a Cartan subalgebra of a semi-simple Lie algebra $\mathfrak{G}$ and ρ is a representation of $\mathfrak{G}$, then $\rho(\mathfrak{H})$ is fully reducible.

If $\mathfrak{G}$ is a Lie algebra with radical $\mathfrak{R}$, then the radical of the commutator subalgebra is easily seen from Levi's decomposition to be in the ideal $[\mathfrak{G}, \mathfrak{R}]$. It follows readily from Lie's theorem on the simultaneous triangularizing of matrices with coefficients in an algebraically closed field that the commutator subalgebra of $\mathrm{ad}(Kg + \mathfrak{R})$ consists of nilpotent elements for $g \varepsilon \mathfrak{G}$. It follows at once that $[\mathfrak{G}, \mathfrak{R}]$ is a nilpotent ideal of $\mathfrak{G}$ and thus the radical of the derived algebra of a Lie algebra is nilpotent. If $\mathfrak{N}$ is a subalgebra of a Lie algebra $\mathfrak{G}$, we mean by an *inner automorphism from* $\mathfrak{N}$ any product of automorphisms of the form $\exp(\mathrm{ad}\, x)$ with x in $\mathfrak{N}$ and $\mathrm{ad}\, x$ nilpotent (there is no difficulty concerning convergence and it can be verified readily that $\exp \mathrm{ad}\, x$ is an automorphism of $\mathfrak{G}$). The inner automorphism $\exp(\mathrm{ad}\, x)$ is a similarity in view of the identity $(\exp(\mathrm{ad}\, x))(y) = (\exp x) \cdot y \cdot (\exp x)^{-1}$.

The Cartan subalgebras of a *solvable* Lie algebra $\mathfrak{G}$ are conjugate under inner automorphisms from $\mathfrak{G}^{\infty}$ (cf. [1], [3]).

2. 3. Groups of unipotent elements. By the replica of an endomorphism x of a linear space is meant an element in the smallest algebraic Lie algebra containing x. A linear Lie algebra is algebraic if and only if it contains all the replicas of its elements (cf. [2]). The replicas of a nilpotent endomorphism n are precisely the multiples of n ([2], p. 159). Thus any linear Lie algebra of nilpotent endomorphisms is algebraic. If n is a nilpotent endomorphism, the connected algebraic Lie group determined by Kn is $\exp Kn$ ([2], p. 159). If u is a unipotent automorphism, its algebraic group hull is $\exp K \log u$ (cf. [2], p. 183) which is connected. Thus every unipotent element of an algebraic group lies in the connected component of the identity.

We shall mean by a rational representation of an algebraic Lie algebra $\mathfrak{G}$ a homomorphism which is the differential at the identity of a rational representation of an algebraic group G whose Lie algebra is $\mathfrak{G}$. Clearly if ρ is a rational representation of $\mathfrak{G}$ into $\mathfrak{H}$, and $\mathfrak{H}_1$ is an algebraic subalgebra of $\mathfrak{H}$, then $\rho^{-1}(\mathfrak{H}_1)$ is algebraic. Conversely, if $\mathfrak{G}$ is algebraic and ρ is rational, then $\rho(\mathfrak{G})$ is algebraic (though this is not always true when the characteristic of the base field is not 0) (cf. [2], pp. 140, 146).

The analogue of these results for algebraic groups is not true; that is, if G is an algebraic group and ρ a rational representation, then in general $\rho(G)$ is not algebraic. However in case G is an algebraic group of unipotent elements, the analogue is true in view of the following special features of such groups.[2]

---

[2] U1, U2, and U3 were proved by the author in the original version of this paper.

U1) An algebraic group of unipotent elements is connected and its Lie algebra consists of nilpotent elements.

U2) Let $\mathfrak{N}$ be a linear Lie algebra of nilpotent elements and let $X_1, \cdots, X_r$ be a base for $\mathfrak{N}$ such that the linear span $\mathfrak{N}_i$ of $X_i, \cdots, X_r$ is an ideal in $\mathfrak{N}_{i-1}$ $(i=2, \cdots, r)$. Such a base always exists. Let N be a connected algebraic group with Lie algebra $\mathfrak{N}$. Then the elements of N can be expressed uniquely as $\exp(t_1X_1) \cdot \exp(t_2X_2) \cdots \exp(t_rX_r)$ with $t_1, \cdots, t_r$ in the ground field. ([3], Ch. V, Sec. 3, Prop. 17).

U3) Let ρ be a rational representation of an algebraic group N whose Lie algebra consists of nilpotent elements. Then $\rho(N)$ is algebraic ([3], Ch. V, Sec. 3, Prop. 15).

It appears from these results that an algebraic group consists of unipotent elements if and only if it is connected and its Lie algebra consists of nilpotent elements; in addition the image of such a group under a rational representation is algebraic.

Section 3. Rational Representations.

This section is devoted to a discussion of the image of fully reducible groups and groups of unipotent elements under a rational representation.

It is easily seen that a normal subgroup of a fully reducible linear group is fully reducible. We now prove a simple but highly useful partial converse.

LEMMA 3.1.[3] *Let N be a normal subgroup of finite index in the linear group G. G is fully reducible if and only if N is fully reducible.*

Proof. In view of the preceding remark, it remains only to prove that if N is fully reducible, then G is fully reducible.

Suppose therefore that N is fully reducible and that the finite set of elements $g_1, \cdots, g_p$ is a complete set of representatives for the cosets of $G \bmod N$. Let G operate on the linear space V and suppose it keeps invariant the subspace W. Let U be a complement to W invariant under N. For

After the paper was submitted for publication, the author learned that these results as well as many other interemediate results of the paper had been obtained independently by C. Chevalley and included in his "Theorie des Groupes de Lie," vol. 3, which was to appear soon. In the present revised version, the author has omitted most proofs of such results, taking advantage of Chevalley's Volume 3 as a reference.

[3] This observation seems to have gone unnoticed in the literature. It was noted independently by C. Chevalley who gives a different proof for it in Prop. 1, Ch. IV, Sec. 5 of [3].

each u in U, let u_i be the unique element of g_iU such that $u_i - u$ is in W $(i = 1, \cdots, p)$. Since N is normal in G, each g_iU is invariant under N and the finite set of subspaces $g_1U, \cdots, g_pU$ is permuted by the operations of G. As a result the finite (unordered) sets of elements $S_u = (u_1, \cdots, u_p)$ are permuted among themselves by the operations of G (according to the rule $gS_u = S_{u_g}$ where u_g is the element of U congruent to $gu \bmod W$). For each u in U, let $\bar{u} = (u_1 + \cdots + u_p)/p$ and let $\bar{U}$ denote the set of all elements $\bar{u}$. Then clearly $\bar{U}$ is a linear subspace complementary to W and invariant under G.

Combining Lemma 3.1 with the fact that a connected algebraic group is fully reducible if and only if its Lie algebra is, we conclude

PROPOSITION 3.1.[4] An algebraic group is fully reducible if and only if its Lie algebra is fully reducible.

PROPOSITION 3.2. *A rational representation of an algebraic group sends unipotent elements into unipotent elements and fully reducible subgroups into fully reducible subgroups.*[5]

Proof. It becomes clear after extending the ground field to its algebraic closure that no generality is lost if we assume the ground field to be algebraically closed. Accordingly we assume the ground field K to be algebraically closed. Let u be a unipotent element of the group and let $n = \log u$, which is in the Lie algebra of the group (cf. Sec. 2.2). Let ρ be the given rational representation and let $d\rho$ be the differential of ρ at the identity. We show that $d\rho(n)$ is nilpotent. We let P denote the field of formal power series in an indeterminate t with coefficients in K, and we extend the ground field to P. Let V_1 be the underlying linear space over K on which n acts, let E_1 denote the space of endomorphisms of V_1, and let $V_1{}^P$, $E_1{}^P$ denote the extensions obtained on extending K to P. Let V_2, E_2, $V_2{}^P$, $E_2{}^P$ be the spaces related similarly to the endomorphism $d\rho(n)$. For any element x of $E_i{}^P$ we denote by $L(x)$ the set of values taken at x by the linear extensions to $E_i{}^P$ of linear functions on E_i $(i = 1, 2)$. $L(x)$ can also be described as the K-linear span of the matrix coefficients of x relative to a base B, where the elements of B are in V_i $(i = 1, 2)$. We consider (cf. [2], Sec. 2, Theorem 9, p. 157) the relation

$$\text{(A)} \qquad \rho(\exp tn) = \exp t(d\rho(n)).$$

[4] Also contained in [3].

[5] The assertion about unipotent elements follows easily from Prop. 16, p. 126 of [3].

Since n is nilpotent, $L(\exp tn)$ consists of polynomials in t with coefficients in K. Since ρ is rational map, $L(\rho(\exp tn))$ consists of rational functions of t. On the other hand, let b be an eigenvalue of $d\rho(n)$. Selecting a base in V_2 so as to make the matrix of $d\rho(n)$ triangular, we see that $L(\exp t(d\rho(n)))$ contains the power series $\exp bt$. In view of the cited identity, $\exp bt$ is a rational function in t, which is absurd if $b \neq 0$. (For if $\exp bt = f/g$ with f and g polynomials in t and $b \neq 0$, we arrive at the contradictions degree $g^2 f =$ degree $g \cdot dg/dt \cdot f$ upon differentiating both sides). Hence $b = 0$ and $d\rho(n)$ is nilpotent.

The proof that ρ sends fully reducible subgroups into fully reducible subgroups runs as follows. In view of Proposition 3.1, it suffices to prove that $d\rho$ carries fully reducible Lie algebras into fully reducible Lie algebras. In view of the fact that a fully reducible Lie algebra is the direct sum of a semi-simple Lie algebra and an abelian algebra of fully reducible elements (cf. [6]) and that moreover every semi-simple linear algebra is fully reducible, ([11]), it suffices to prove that $d\rho$ carries fully reducible elements into fully reducible elements. Suppose then that X is a fully reducible endomorphism of V_1. Let $s + n$ be a Jordan sum decomposition of $Y = d\rho(X)$. Let $a_1, \cdots, a_n$ be the eigenvalues of X and $b_1, \cdots, b_n$ the eigenvalues of s. Upon diagonalizing X it is seen that $L(\exp tX)$ is the linear span of $\exp a_1 t, \cdots, \exp a_n t$. Suppose now $n \neq 0$. Upon putting Y into Jordan normal form, it is seen that $L(\exp tY)$ contains a term $t(\exp b_i t)$. Equation (A) above implies that $t(\exp b_i t)$ is a rational combination of $\exp a_1 t, \cdots$, $\exp a_n t$—which is absurd (cf. [2], p. 151). The proof is now complete. Our proof has also established

PROPOSITION 3.3. *A rational representation of an algebraic Lie algebra carries fully reducible subalgebras into fully reducible subalgebras and nilpotent elements into nilpotent elements.*

There is a very elegant proof due to M. Schiffer of the First Main Theorem of the Theory of Invariants for linear groups all of whose "powers" or tensor representations are fully reducible (cf. [9], p. 300). Inasmuch as tensor representations are rational, we have as a consequence of Proposition 3.2.

THEOREM. *The First Main Theorem of the Theory of Invariants is valid for fully reducible groups.*

Section 4. Conjugacy of Fully Reducible Subalgebras.

Throughout this section, $\mathfrak{G}$ denotes a linear Lie algebra, $\mathfrak{R}$ its radical, and $\mathfrak{N}$ the set of nilpotent endomorphisms in $\mathfrak{R}$.

4. 1. If G is an algebraic Lie algebra, it is known that it contains a semi-simple Lie subalgebra $\mathfrak{L}$ and an abelian subalgebra $\mathfrak{A}$ of fully reducible endomorphisms such that $G = \mathfrak{L} + \mathfrak{A} + \mathfrak{N}$, $[\mathfrak{L}, \mathfrak{A}] = 0$, and $\mathfrak{R} = \mathfrak{A} + \mathfrak{N}$ (semi-direct) (cf. [4]). Inasmuch as $\mathfrak{L}$ is fully reducible ([11]), it is easily seen that $\mathfrak{L} + \mathfrak{A}$ is fully reducible. Now $\mathfrak{N}$ can clearly be characterized as the maximal ideal of nilpotent endomorphisms (by applying in succession Engel's theorem and Lie's theorem on triangularizing solvable linear Lie algebras). Consequently a criterion that an algebraic Lie algebra be fully reducible is that it contain non non-zero ideal of nilpotent elements. It may be noted that $\mathfrak{N}$ is the totality of nilpotent endomorphisms in the radical.

4. 2. If A, B are subspaces of a linear space E with $A \subset B$, then the totality of endomorphisms of E which send B into A is an algebraic Lie algebra ([2]). Coupling this fact with the fact that the adjoint representation of an algebraic Lie algebra is rational, we see: If $\mathfrak{G} \supset \mathfrak{B} \supset \mathfrak{A}$ are linear Lie algebras with $[\mathfrak{G}, \mathfrak{B}] \subset \mathfrak{A}$, then $[\mathfrak{G}^*, \mathfrak{B}] \subset \mathfrak{A}$, where $\mathfrak{G}^*$ denotes the smallest algebraic Lie algebra containing $\mathfrak{G}$.

If $\mathfrak{G} \supset \mathfrak{A}$ are linear Lie algebras, the *idealizer* of $\mathfrak{A}$ in $\mathfrak{G}$ is defined as the totality of elements x in G with $[x, \mathfrak{A}] \subset \mathfrak{A}$. If $\mathfrak{G}$ is algebraic, then by the foregoing the idealizer of $\mathfrak{A}$ in $\mathfrak{G}$ is algebraic. By simply "the idealizer of $\mathfrak{A}$" is meant the idealizer of $\mathfrak{A}$ in the totality of endomorphisms.

The radical of an algebraic Lie algebra is algebraic and so also is any subalgebra of nilpotent endomorphisms (cf. [3]).

4. 3. An ideal $\mathfrak{B}$ of a fully reducible linear Lie algebra $\mathfrak{G}$ is fully reducible. For letting $\mathfrak{B}^*$ and $\mathfrak{G}^*$ denote the smallest algebraic Lie algebras containing $\mathfrak{B}$ and $\mathfrak{G}$, it is seen in turn that

$$[\mathfrak{B}, \mathfrak{G}^*] \subset [\mathfrak{B}, \mathfrak{G}] \subset \mathfrak{B}, \qquad [\mathfrak{B}^*, \mathfrak{G}^*] \subset [\mathfrak{B}, \mathfrak{G}^*] \subset \mathfrak{B} \subset \mathfrak{B}^*,$$

and therefore $\mathfrak{B}^*$ is an ideal of $\mathfrak{G}^*$. It is also seen that $\mathfrak{G}^*$ is fully reducible and $\mathfrak{B}$ is fully reducible if and only if $\mathfrak{B}^*$ is. Now the maximum ideal of nilpotent elements of $\mathfrak{B}^*$ is an ideal in $\mathfrak{G}^*$ and hence zero. Therefore $\mathfrak{B}^*$ is fully reducible and thus $\mathfrak{B}$ is too.

Any abelian ad-reductive subalgebra of a Lie algebra can be extended to a Cartan subalgebra (cf. [1] or [3]).

4.4. Suppose $\mathfrak{M}$ and $\mathfrak{N}$ are subalgebras which span the linear Lie algebra $\mathfrak{G}$. Suppose in addition that $\mathfrak{N}$ is an ideal and algebraic. Then $\mathfrak{G}$ is algebraic if and only if $\mathfrak{M}$ is.

Proof. Let ρ be a rational representation of $\mathfrak{F}$, the idealizer of $\mathfrak{N}$, whose kernel is $\mathfrak{N}$. (That such a representation exists is stated in [2] and proved in [3]). Then $\mathfrak{G}=\rho^{-1}(\rho(\mathfrak{M}))$ is algebraic according as $\mathfrak{M}$ is.

LEMMA 4.1. *Let ρ be a representation of a solvable Lie algebra $\mathfrak{G}$. Then any two maximal ρ-reductive subalgebras of $\mathfrak{G}$ are conjugate under an inner automorphism from $\mathfrak{G}^{\infty}$. If ρ is an isomorphism, a ρ-reductive subalgebra is abelian.*

Proof. Since $\rho(\mathfrak{G}^{\infty})=\rho(\mathfrak{G})^{\infty}$, it suffices to prove the theorem for a solvable linear Lie algebra $\mathfrak{G}$ (with ρ the identity map). Let $\mathfrak{A}_i$ $(i=1,2)$ be a maximal fully reducible subalgebra of $\mathfrak{G}$. $[\mathfrak{A}_i,\mathfrak{A}_i]$ consists of elements which are fully reducible and nilpotent (by Lie's triangularizing Theorem) and is thus (0); that is $\mathfrak{A}_i$ is abelian. Since $\mathfrak{A}_i$ is an abelian set of ad-reductive elements, it is an ad-reductive subalgebra. By 4.3 it can be extended to a Cartan subalgebra $\mathfrak{H}_i$ $(i=1,2)$. $\mathfrak{G}$ being solvable, there is an inner automorphism x from $\mathfrak{G}^{\infty}$ which sends $\mathfrak{H}_1$ into $\mathfrak{H}_2$ (cf. [1]). Now $\mathfrak{H}_i$ being a nilpotent linear Lie algebra, the fully reducible elements in it are simultaneously ad-reductive and ad-nilpotent and hence central. Hence they form a subalgebra—in fact $\mathfrak{A}_i$. The automorphism x which carries $\mathfrak{H}_1$ into $\mathfrak{H}_2$ carries $\mathfrak{A}_1$ into $\mathfrak{A}_2$ since it sends fully reducible elements into fully reducible elements (for it is a similarity). Proof of the lemma is now complete.

LEMMA 4.2. *Let $\mathfrak{G}$ be a linear Lie algebra, $\mathfrak{N}$ an ideal of nilpotent elements in $\mathfrak{G}$, and $\mathfrak{M}$, $\mathfrak{M}_1$ fully reducible subalgebras of $\mathfrak{G}$ with $\mathfrak{M}_1\subset\mathfrak{M}+\mathfrak{N}$. Then $\mathfrak{M}\cap(\mathfrak{M}_1+\mathfrak{N})$ is fully reducible.*

Proof. Let $\mathfrak{F}$ be the idealizer of $\mathfrak{N}$. It is algebraic and contains $\mathfrak{G}$. Let ρ be a rational representation of $\mathfrak{F}$ with kernel $\mathfrak{N}$, and let ρ_0 be the restriction of ρ to $\mathfrak{M}^*$, the smallest algebraic Lie algebra containing $\mathfrak{M}$. Since $\mathfrak{M}^*$ is fully reducible and $\mathfrak{M}^*\cap\mathfrak{N}$ is an ideal of nilpotent elements in $\mathfrak{M}^*$, it is zero (by Remark 4.1). Hence ρ is one-to-one on $\mathfrak{M}^*$. Furthermore, since rational representations carry algebraic Lie algebras into algebraic Lie algebras and since inverse images under rational representations of algebraic

Lie algebras are algebraic, we have $\rho_0(\mathfrak{M}^*_2) = \rho_0(\mathfrak{M}_2)^*$ where $\mathfrak{M}_2$ is the subalgebra $\mathfrak{M} \cap (\mathfrak{M}_1 + \mathfrak{N})$ and the asterisk denotes closure in the sense of smallest containing algebraic Lie algebra. Now it is clear that $\rho_0(\mathfrak{M}_2) = \rho(\mathfrak{M}_2) = \rho(\mathfrak{M}_1)$ and hence $\rho_0(\mathfrak{M}_2)$ is fully reducible (Proposition 3.3). Consequently, $\rho_0(\mathfrak{M}^*_2) = \rho_0(\mathfrak{M}_2)^*$ is fully reducible and hence $\rho_0(\mathfrak{M}_2)^*$ contains no non-zero ideal of nilpotent elements by Remark 4.1. Since ρ_0 takes nilpotent elements into nilpotent elements and it is one-to-one on $\mathfrak{M}^*_2$, we conclude that $\mathfrak{M}^*_2$ contains no ideal of nilpotent elements. It follows that $\mathfrak{M}^*_2$ is fully reducible and hence $\mathfrak{M}_2$ is fully reducible.

THEOREM 4.1. *Any two maximal fully reducible subalgebras of a linear Lie algebra $\mathfrak{G}$ are conjugate under an inner automorphism from the radical of $[\mathfrak{G}, \mathfrak{G}]$.*

Proof. Let $\mathfrak{G}^*$ be the smallest algebraic Lie algebra containing $\mathfrak{G}$ and let $\mathfrak{M}$ be a fully reducible subalgebra of $\mathfrak{G}^*$. Then $\mathfrak{M} \cap \mathfrak{G}$ is an ideal in $\mathfrak{M}$ and hence fully reducible (Remark 4.1). Again since $[\mathfrak{G}^*, \mathfrak{G}^*] = [\mathfrak{G}, \mathfrak{G}]$, and moreover some maximal fully reducible subalgebra of $\mathfrak{G}^*$ contains a maximal fully reducible subalgebra of $\mathfrak{G}$, Theorem 4.1 would be valid for $\mathfrak{G}$ if it were valid for algebraic Lie algebras.

Thus we may adopt the additional hypothesis that $\mathfrak{G}$ is algebraic without loss of generality. We proceed by induction on the dimension of $\mathfrak{G}$.

Let $\mathfrak{N}$ denote the ideal of nilpotent elements in the radical of the algebraic Lie algebra $\mathfrak{G}$, and let $\mathfrak{M}_1$ be a fully reducible subalgebra such that $\mathfrak{G} = \mathfrak{M}_1 + \mathfrak{N}$ (cf. 4.1). $\mathfrak{M}_1$ is a maximal fully reducible subalgebra, since any fully reducible algebra $\mathfrak{M}_0$ containing it must intersect $\mathfrak{N}$ in an ideal of $\mathfrak{M}_0$, and hence (Remark 4.1) in zero. Furthermore, the radical of $[\mathfrak{G}, \mathfrak{G}]$ is $[\mathfrak{M}_1, \mathfrak{N}]$ (Remark 4.1). Let now $\mathfrak{M}_2$ be a maximal fully reducible subalgebra of $\mathfrak{G}$. Then $\mathfrak{M}_1 \cap (\mathfrak{M}_2 + \mathfrak{N})$ is fully reducible by Lemma 4.2, indeed maximal fully reducible in $\mathfrak{M}_2 + \mathfrak{N}$ since it is complementary to $\mathfrak{N}$ in $\mathfrak{M}_2 + \mathfrak{N}$. Furthermore $\mathfrak{M}_2$ is algebraic and hence $\mathfrak{M}_2 + \mathfrak{N}$ is algebraic. Hence we must have $\mathfrak{M}_2 + \mathfrak{N} = \mathfrak{G}$; otherwise on applying the induction assumption we would be led to the false conclusion that $\mathfrak{M}_1 \cap (\mathfrak{M}_2 + \mathfrak{N})$ is conjugate to a maximal fully reducible subalgebra of $\mathfrak{G}$ and hence is $\mathfrak{M}_1$. Now let $\mathfrak{L}_i + \mathfrak{A}_i$ be a Levi-decomposition for the fully reducible subalgebra $\mathfrak{M}_i$ $(i = 1, 2)$, $\mathfrak{L}_i$ being semi-simple. Then it is easily seen that $\mathfrak{L}_i$ is a maximal semi-simple subalgebra of $\mathfrak{G}$. As is well-known, the maximal semi-simple subalgebras of a Lie algebra (over a field of characteristic zero) are conjugate under an inner automorphism from the radical of the commutator subalgebra (cf. [5]). Carrying $\mathfrak{L}_2$

into $\mathfrak{L}_1$ by such an automorphism, we see that no generality is lost in assuming $\mathfrak{L}_1 = \mathfrak{L}_2$. Again, since $\mathfrak{M}_1$ and $\mathfrak{M}_2$ are each contained in the idealizer of $\mathfrak{L}_1$ in $\mathfrak{G}$, no generality is lost in assuming that $\mathfrak{L}_1$ is an ideal in $\mathfrak{G}$. In this case, $\mathfrak{G}$ is the direct sum of $\mathfrak{L}_1$ and its radical $\mathfrak{R}$. Now $\mathfrak{M}_i$ being maximal fully reducible in $\mathfrak{G}$, and fully reducible subalgebras of $\mathfrak{R}$ being abelian and commutative with $\mathfrak{L}_1$, it is easily seen that $\mathfrak{M}_i \cap \mathfrak{R}$ is a maximal fully reducible subalgebra of $\mathfrak{R}$. Applying Lemma 4.1 to $\mathfrak{R}$, we can get an inner automorphism from $\mathfrak{R} \cap [\mathfrak{G}, \mathfrak{G}]$ which sends $\mathfrak{M}_2$ into $\mathfrak{M}_1$. The proof of the theorem is now complete.

COROLLARY 4.1. *Let $\mathfrak{G}$ be a linear Lie algebra with radical $\mathfrak{R}$. Assume $\mathfrak{A}$ is a maximal fully reducible abelian subalgebra of $\mathfrak{G}$. Then there is a maximal semi-simple subalgebra $\mathfrak{L}$ of $\mathfrak{G}$ such that $\mathfrak{A} = \mathfrak{A} \cap \mathfrak{L} + \mathfrak{A} \cap \mathfrak{R}$ with (1) $\mathfrak{A} \cap \mathfrak{L}$ a Cartan subalgebra of $\mathfrak{L}$, (2) $\mathfrak{A} \cap \mathfrak{R}$ a maximal fully reducible subalgebra of $\mathfrak{R}$ and (3) $[\mathfrak{L}, \mathfrak{A} \cap \mathfrak{R}] = 0$.*

Proof. Let $\mathfrak{B}$ be a maximal fully reducible subalgebra of $\mathfrak{R}$. Since $\mathfrak{R}$ is an ideal in $\mathfrak{G}$, we have $\mathfrak{G} = Z(\mathfrak{B}) + \mathfrak{R}$ where $Z(\mathfrak{B})$ is the centralizer of $\mathfrak{B}$. Letting $\mathfrak{L}$ be the semi-simple component in a Levi-decomposition for $Z(\mathfrak{B})$, we see that $\mathfrak{G} = \mathfrak{L} + \mathfrak{R}$ and $\mathfrak{L}$ is a maximal semi-simple subalgebra of $\mathfrak{G}$. Moreover, $\mathfrak{L} + \mathfrak{B}$ is readily seen to be maximal fully reducible in $\mathfrak{G}$. Any maximal fully reducible abelian subalgebra of $\mathfrak{L} + \mathfrak{B}$ clearly has the properties (1), (2), (3) of the Corollary. Inasmuch as the maximal fully reducible subalgebras are conjugate under inner automorphisms, any maximal fully reducible abelian subalgebra has the desired properties.

Another consequence, which follows directly from the theorem is

COROLLARY 4.2. *Any two maximal ad-reductive subalgebras of a Lie algebra are conjugate under an inner automorphism from the radical of the commutator subalgebra.*

Section 5. Invariant Subalgebras of a Fully Reducible Group of Automorphisms.

LEMMA 5.1. *Let $\mathfrak{M}$ be a subalgebra and $\mathfrak{N}$ an abelian ideal of a Lie algebra $\mathfrak{G}$, and let Γ be a group of automorphisms of $\mathfrak{G}$. Assume*

1) $\mathfrak{G} = \mathfrak{M} + \mathfrak{N}$ *semi-directly;*

2) *for each automorphism T of Γ there is an element n in $\mathfrak{N}$ such that $T(\mathfrak{M}) = \exp \operatorname{ad} n(\mathfrak{M})$;*

3) Γ *keep* $\mathfrak{N}$ *invariant.*

Then Γ *keeps invariant some image of* $\mathfrak{M}$ *by an inner automorphism.*

Proof. Let A denote the totality of linear subspaces of $\mathfrak{G}$ that are complementary to $\mathfrak{N}$. We regard A as an affine space in the natural way; viz, if we regard $\mathfrak{M}$ as the "origin" or "reference point" of A, then the linear maps of $\mathfrak{M}$ into $\mathfrak{N}$ can be identified with the associated vector space V of A under the correspondence $\mathfrak{L} - \mathfrak{M} \rightarrow \phi(\mathfrak{L})$ where $\mathfrak{L} \varepsilon A$ and $\phi(\mathfrak{L})(m)$ is the element p in $\mathfrak{N}$ such that $m + p \varepsilon \mathfrak{L}$.

The totality S of images of $\mathfrak{M}$ by inner automorphisms forms an affine subspace of A. For $\mathfrak{N}$ being an abelian ideal, $(\operatorname{ad} n)^2 = 0$ for n in N and thus $\exp \operatorname{ad} n(m) = m + [n, m]$; consequently the family S is identified with the linear family of maps $m \rightarrow [n, m]$ (n varying over $\mathfrak{N}$) of $\mathfrak{M}$ into $\mathfrak{N}$.

Now the fully reducible group of automorphisms Γ of $\mathfrak{G}$ induces on A a group of affine transformations and has in A a fixed point $\mathfrak{M}_0$. Selecting $\mathfrak{M}_0$ as origin in A, we obtain a representation of the group of automorphisms of $\mathfrak{G}$ which keep $\mathfrak{M}_0$ invariant on the vector space V associated with A and this representation is easily seen to be rational. Since a rational representation takes fully reducible groups into fully reducible groups, it follows that Γ^* the image of Γ under this representation is fully reducible. Furthermore the operations of Γ on A keep invariant the affine subspace S. Hence Γ^* keeps invariant the linear subspace $U = S - \mathfrak{M}$ of V. Being fully reducible, Γ^* keeps invariant a subspace W complementary to U. Hence the operations of Γ on A keep invariant the affine subspace $\mathfrak{M}_0 + W$, which intersects $S = \mathfrak{M} + U$ in but a single point—call it $\mathfrak{M}_1$. Obviously $\mathfrak{M}_1$ is a fixed point of Γ^* and correspondingly the subalgebra $\mathfrak{M}_1$ is invariant under Γ and is the image of $\mathfrak{M}$ under an inner automorphism from $\mathfrak{N}$. Proof of the lemma is now complete.

Note. The analogue of Lemma 5.1 for associative algebras is true and proof is exactly the same.

DEFINITION. *Let* $\mathfrak{G}$ *be a linear Lie algebra, let* $\mathfrak{N}$ *be an ideal of* $\mathfrak{G}$, *and let* T *be a similarity of* $\mathfrak{G}$ *which keeps* $\mathfrak{N}$ *invariant. Let* ρ *be a rational representation of* $\mathfrak{G}$ *with kernel* $\mathfrak{N}$, *and let* T_1 *denote the automorphism of* $\rho(G)$ *induced by* T. *Then* T_1 *is called an s-automorphism of* $\rho(G)$.

Clearly an s-automorphism of a linear Lie algebra carries maximal fully reducible subalgebras into maximal fully reducible subalgebras and nilpotent elements into nilpotent elements.

THEOREM 5.1. *A pre-fully reducible group of similarities of a linear Lie algebra keeps invariant a maximal fully reducible subalgebra.*

Proof. Let $\mathfrak{G}$ be the linear Lie algebra and $\mathfrak{G}^*$ the smallest algebraic Lie algebra containing $\mathfrak{G}$.

We remark first that any maximal fully reducible subalgebra of $\mathfrak{G}^*$ intersects $\mathfrak{G}$ in a maximal fully reducible subalgebra of $\mathfrak{G}$. For obviously some maximal fully reducible subalgebra $\mathfrak{M}^*$ of $\mathfrak{G}^*$ includes a maximal fully reducible subalgebra $\mathfrak{M}$ of $\mathfrak{G}$. Since $[\mathfrak{G}^*, \mathfrak{G}^*] = [\mathfrak{G}, \mathfrak{G}]$, $\mathfrak{M}^* \cap \mathfrak{G}$ is an ideal in $\mathfrak{M}^*$ and fully reducible by Remark 4.3. Hence $\mathfrak{M}^* \cap \mathfrak{G} = \mathfrak{M}$. Inasmuch as any maximal fully reducible subalgebra of $\mathfrak{G}^*$ is conjugate to $\mathfrak{M}^*$ by an inner automorphism from $[\mathfrak{G}^*, \mathfrak{G}^*] \subset \mathfrak{G}$, it follows at once that *every* maximal fully reducible subalgebra of $\mathfrak{G}^*$ intersects $\mathfrak{G}$ in a maximal fully reducible subalgebra of $\mathfrak{G}$.

Next we note that any similarity T keeping $\mathfrak{G}$ invariant also keeps $\mathfrak{G}^*$ invariant. For $T(\mathfrak{G}^*) \cap \mathfrak{G}^*$ is algebraic and includes $\mathfrak{G}$. Hence it includes $\mathfrak{G}^*$; consequently $T(\mathfrak{G}^*) = \mathfrak{G}^*$.

In view of the preceding observation, we can assume without loss of generality that $\mathfrak{G}$ is an algebraic Lie algebra. Thus the theorem will be established when we will have proved: a fully reducible group of s-automorphisms of an algebraic Lie algebra $\mathfrak{G}$ keeps invariant a maximal fully reducible subalgebra of $\mathfrak{G}$.

We decompose $\mathfrak{G}$ into $\mathfrak{M} + \mathfrak{N}$ where $\mathfrak{M}$ is a maximal fully reducible subalgebra and $\mathfrak{N}$ is the maximum ideal of nilpotent elements. We prove the above assertion by induction on $r = \dim[\mathfrak{N}, \mathfrak{N}]$.

If $\dim[\mathfrak{N}, \mathfrak{N}] = 0$, the assertion follows immediately from Lemma 5.1. Assume now that $\dim[\mathfrak{N}, \mathfrak{N}] > 0$ and that the induction hypothesis holds for $r < \dim[\mathfrak{N}, \mathfrak{N}]$. $[\mathfrak{N}, \mathfrak{N}]$ is an algebraic Lie algebra. Hence there is a rational representation ρ of $\mathfrak{G}$ whose kernel is $[\mathfrak{N}, \mathfrak{N}]$. Applying the induction hypothesis to the induced fully reducible group of s-automorphisms of $\rho(\mathfrak{G})$, we are reduced to verifying the hypothesis for $\mathfrak{M}_1 + [\mathfrak{N}, \mathfrak{N}]$ where $\mathfrak{M}_1$ is the image of $\mathfrak{M}$ by an inner automorphism from $\mathfrak{N}$. Since an inner automorphism is a similarity, $\mathfrak{M}_1$ is a maximal fully reducible subalgebra of $\mathfrak{G}$ and hence algebraic. Consequently $\mathfrak{M}_1 + [\mathfrak{N}, \mathfrak{N}]$ is algebraic by 4.4. Furthermore $\mathfrak{N}$ is solvable by Engel's theorem so that the dimension of the commutator subalgebra of $[\mathfrak{N}, \mathfrak{N}]$ is smaller than $\dim[\mathfrak{N}, \mathfrak{N}]$. We can thus apply the induction hypothesis to $\mathfrak{M}_1 + [\mathfrak{N}, \mathfrak{N}]$ and thereby complete the proof of Theorem 5.1.

COROLLARY 5.1. *A fully reducible group of automorphisms of a Lie algebra keeps invariant a maximal ad-reductive subalgebra.*

Proof. Once we note that the center of a Lie algebra belongs to every maximal ad-reductive subalgebra, this corollary follows directly from Theorem 5.1.

COROLLARY 5.2. *A fully reducible group of automorphisms of a Lie algebra keeps invariant a maximal semi-simple subalgebra.*

Proof. Follows directly from Corollary 5.1 and the observation that an ad-reductive subalgebra has a unique maximum semi-simple subalgebra—viz., its commutator subalgebra.

THEOREM 5.2. *A fully reducible group of automorphisms of a solvable Lie algebra keeps invariant a Cartan subalgebra.*

Proof. Let $\mathfrak{G}$ be the solvable Lie algebra and Γ the fully reducible group of automorphisms. The operation $\operatorname{ad} X \rightarrow \operatorname{ad} c(X)$, $c \varepsilon \Gamma$, of c on $\operatorname{ad} \mathfrak{G}$ is a similarity since $\operatorname{ad} c(X) = c(\operatorname{ad} X)c^{-1}$. Inasmuch as any Cartan subalgebra of $\mathfrak{G}$ is the inverse image of a Cartan subalgebra of $\operatorname{ad} \mathfrak{G}$ (and conversely), it suffices to prove

(A) any pre-fully reducible group of similarities of a solvable linear Lie algebra keeps invariant a Cartan subalgebra.

Now it is not difficult to see that any Cartan subalgebra of a linear Lie algebra $\mathfrak{L}$ is the intersection with $\mathfrak{L}$ of a Cartan subalgebra of the algebraic hull $\mathfrak{L}^*$ and conversely (cf. [3]). Consequently it suffices to prove (A) under the additional hypothesis that the solvable linear Lie algebra is algebraic.

We can suppose accordingly that Γ is a pre-fully reducible group of similarities of a solvable algebraic Lie algebra $\mathfrak{G}$. By Corollary 5.1 Γ keeps invariant a maximal fully reducible subalgebra $\mathfrak{M}$ of $\mathfrak{G}$. Now any algebraic Lie algebra obviously contains some regular fully reducible element x and the nilspace of $(\operatorname{ad} x)_{\mathfrak{G}}$ is a Cartan subalgebra. Furthermore any fully reducible subalgebra $\mathfrak{F}$ of a solvable linear Lie algebra is abelian—for its ideal $[\mathfrak{F}, \mathfrak{F}]$ being fully reducible and consisting of nilpotent elements is zero. Thus any maximal ad-reductive subalgebra $\mathfrak{M}_x$ of $\mathfrak{G}$ which contains x extends to a unique Cartan subalgebra—the nilspace of $(\operatorname{ad} \mathfrak{M}_x)_{\mathfrak{G}}$ or the centralizer of $\mathfrak{M}_x$. Since all maximal fully reducible subalgebras of $\mathfrak{G}$ are conjugate, the centralizer of any maximal fully reducible subalgebra of $\mathfrak{G}$ is a Cartan

subalgebra. In particular, the centralizer of $\mathfrak{M}$ is a Cartan subalgebra invariant under Γ. Proof of the theorem is now complete.

By paralleling the proof of Theorem 5.1 we can prove

THEOREM. *A fully reducible group of automorphisms of an associative or Jordan algebra keeps invariant a maximal fully reducible subalgebra.*

The pertinent facts about Jordan algebras required for the above proof can be found in the paper by N. Jacobson in *Transactions of the American Mathematical Society*, vol. 70, (1951), p. 528.

Section 6. A Decomposition of Algebraic Groups.

Let G be an algebraic group. We denote by $\mathfrak{M}$ a maximal fully reducible subalgebra of its Lie algebra $\mathfrak{G}$, and by $\mathfrak{N}$ the ideal of nilpotent elements in its radical. We denote by M^A the totality of elements x in G such that $\operatorname{Ad} x$ keeps $\mathfrak{M}$ invariant, that is $x\mathfrak{M}x^{-1} = \mathfrak{M}$. Because of its maximal character, $\mathfrak{M}$ is algebraic, and because $\mathfrak{N}$ consists of nilpotent elements only, it, too, is algebraic ([2], pp. 181 and 183). We denote by N the connected group with Lie algebra $\mathfrak{N}$. N is normal in G since $\mathfrak{N}$ is invariant under all similarities from G.

LEMMA 6.1. *Let G be an algebraic group, M^A and N the subgroups of G defined above. Then $G = M^A \cdot N$.*

Proof. Since an algebraic group contains the Jordan product components of each of its elements, it suffices to prove that $M^A \cdot N$ contains each fully reducible and each unipotent element of G.

Suppose that u is unipotent. Then $u = \exp X$ with $\log u = X$ a nilpotent element in the Lie algebra $\mathfrak{G}$. Inasmuch as $KX + \mathfrak{N}$ is a solvable Lie algebra, by Lie's theorem on simultaneous triangularizing it consists of nilpotent elements. Let X_1 be a non-zero element of $(KX + \mathfrak{N}) \cap \mathfrak{M}$ if such exists, otherwise zero. Choosing a suitable base for $\mathfrak{N}$, we can apply U2 of 2.3 to conclude that $u = \exp(t_1X_1) \cdot u'$ with u' in N. Now $\exp t_1X_1$ belongs to the connected algebraic group with Lie algebra KX_1 and hence belongs to the connected algebraic group corresponding to $\mathfrak{M}$; hence $\operatorname{Ad}(\exp t_1X_1)(\mathfrak{M}) = \mathfrak{M}$ and $\exp t_1X_1$ is in M^A. It follows that u is in $M^A \cdot N$.

Suppose on the other hand that s is a fully reducible element of G. By Theorems 5.1 and 4.1 there exists an element u in N such that $s(uMu^{-1})s^{-1} = u\mathfrak{M}u^{-1}$. Hence $(u^{-1}su)\mathfrak{M}(u^{-1}su)^{-1} = \mathfrak{M}$ and $u^{-1}su = m$ is in

M^A. But $\mathfrak{N}$ being a normal subgroup of G, sus^{-1} is in N and hence $v = u^{-1}(sus^{-1})$ is in N. Since $vs = u^{-1}su = m$, we have $s = mv^{-1}$ is in $M^A \cdot N$. Proof of Lemma 6.1 is now complete.

THEOREM 6.1. *Let G be an algebraic group, $\mathfrak{M}$ a maximal fully reducible subalgebra of its Lie algebra $\mathfrak{G}$, and $\mathfrak{N}$ the ideal of nilpotent elements in the radical of $\mathfrak{G}$. There is a fully reducible algebraic group M with Lie algebra $\mathfrak{M}$ such that $G = M \cdot N$ (semi-direct), N being the connected algebraic group with Lie algebra $\mathfrak{N}$.*

Proof. By Lemma 6.1, $G = M^A \cdot N$ (not necessarily semi-direct). It is easily seen from the definition of M^A that it is an algebraic group whose Lie algebra $\mathfrak{G}_1$ contains $\mathfrak{M}$.

We first note that it is sufficient to prove Theorem 6.1 for the group M^A. For clearly $\mathfrak{G}_1 = \mathfrak{M} + (G_1 \cap \mathfrak{N})$ and $\mathfrak{G}_1 \cap \mathfrak{N}$ is the ideal of nilpotent element in the radical of $\mathfrak{G}_1$. If now $M^A = M \cdot N_1$ (semi-direct) where M and N_1 are algebraic groups with Lie algebras $\mathfrak{M}$ and $\mathfrak{G}_1 \cap \mathfrak{N}$ respectively, then $G = MN_1N = MN$. Moreover $M \cap N$, being an algebraic group with Lie algebra $\mathfrak{M} \cap \mathfrak{N} = (0)$, is a finite subgroup of N; being fully reducible it reduces to the identity element. Thus $G = M \cdot N$ (semi-direct).

Hence we lose no generality in assuming $G = M^A$. In this case $\mathfrak{M}$ is an ideal of $\mathfrak{G}$, and $\mathfrak{M}$ being ad-reductive (cf. Prop. 3.2) $\mathfrak{G} = \mathfrak{M} + \mathfrak{N}$ (direct). We proceed by induction on the dimension of $\mathfrak{G}$.

We let $\mathfrak{N}_1$ denote the center of $\mathfrak{N}$, and we denote by N_1 and M_0 the connected algebraic groups corresponding to $\mathfrak{N}_1$ and $\mathfrak{M}$ respectively. Clearly $\mathfrak{N}_1$ and $\mathfrak{M}$ are invariant under similarities from G, and hence N_1 and M_0 are normal subgroups of G. Furthermore, if $\mathfrak{N} \neq 0$ and $\mathfrak{G} \neq \mathfrak{N}_1$, there is a normal algebraic subgroup F of positive dimension such that the dimension of the subalgebra $\mathfrak{F} + \mathfrak{M}$ is smaller than $\dim \mathfrak{G}$ ($\mathfrak{F}$ being the Lie algebra of F)—namely $F = M_0$ or, if M_0 is zero dimensional $F = N_1$.

Case 1. $\mathfrak{N} = 0$. Here $\mathfrak{G} = \mathfrak{M}$ is fully reducible and hence G is fully reducible (Prop. 3.1).

Case 2. $\mathfrak{N} \neq 0$, $\mathfrak{G}$ is not the center of N. Select the subgroup F described above. Selecting a rational representation ρ with kernel F we know by Prop. 3.2 that $\mathfrak{M}' = d\rho(\mathfrak{M})$ is fully reducible, and $\mathfrak{N}' = d\rho(\mathfrak{N})$ is composed of nilpotent elements. By the induction hypothesis $G' = M' \cdot N'$ (semi-direct) where G' is the algebraic group hull of $\rho(G)$, M' is an algebraic subgroup with Lie algebra $\mathfrak{M}'$ and N' is the connected algebraic group with

Lie algebra $\mathfrak{N}'$. By U2 of 2.3 $\rho(N) = N'$. Hence $G = \rho^{-1}(M') \cdot N$. Applying the induction assumption to $\rho^{-1}(M')$, we obtain the desired fully reducible subgroup with Lie algebra $\mathfrak{M}$.

Case 3. $\mathfrak{G}$ is the center of $\mathfrak{N}$. Here $\mathfrak{G} = \mathfrak{N}$ is abelian, and G is a finite extension of N. Furthermore by U2 of 2.3 and the formula $\exp X \exp Y = \exp(X + Y)$ if $[X, Y] = 0$, we see that $X \rightarrow \exp X$ is an isomorphism of the additive group of the linear space $\mathfrak{N}$ with the multiplicative group N. Our theorem will therefore follow from the following: Let N be a vector group over a field K of characteristic zero (i. e. the additive group of a linear space over K—possibly infinite dimensional); then any finite extension of N splits. This in turn is the group extension interpretation of the now well known result: the two dimensional cohomology group of a finite group in an indefinitely divisible abelian group vanishes. (Indeed the cohomology groups vanish in all positive dimensions.)

Section 7. Conjugacy of Fully Reducible Subgroups.

THEOREM 7.1. *Let G be an algebraic group, $\mathfrak{N}$ the set of nilpotent elements in the radical of its Lie algebra, and N the connected algebraic group with Lie algebra $\mathfrak{N}$. Then $G = M \cdot N$ (semi-direct) with M a maximal fully reducible subgroup. Furthermore any two maximal fully reducible subgroups of G are conjugate under an inner automorphism from $\mathfrak{N}$.*

Proof. If M is a fully reducible subgroup of the algebraic group G with $G = M \cdot N$, then M is algebraic and *maximal* fully reducible and $G = M \cdot N$ (semi-direct). For let $\bar{M}$ be a maximal fully reducible subgroup containing M. $\bar{M}$ is clearly algebraic and its Lie algebra $\mathfrak{M}$ is fully reducible. Since $\mathfrak{M} \cap \mathfrak{N}$ is an ideal of nilpotent element in the radical of $\mathfrak{M}$, by Remark 4.1 it follows that $\mathfrak{M} \cap \mathfrak{N} = 0$. Since $\bar{M} \cap N$ is therefore a finite algebraic group, it consists of fully reducible unipotent elements and hence reduces to the identity. Thus $G = \bar{M} \cdot N$ (semi-direct) and $M = \bar{M}$.

We now prove the second part of the theorem by induction on the dimension of $\mathfrak{G}$, the Lie algebra of G. Let M_1 be a maximal fully reducible subgroup of G and $\mathfrak{M}_1$ its Lie algebra. Then $\mathfrak{F} = \mathfrak{M}_1 + \mathfrak{N}$ is algebraic (Sec. 4, Remark 4.4). Let F_0 denote the corresponding connected algebraic group and let F be the finite (algebraic) extension $M_1 F_0$. Now $G = M \cdot N$ implies that $F = (F \cap M) \cdot N$. Since $\mathfrak{F} \cap \mathfrak{M} = \mathfrak{M} \cap (\mathfrak{M}_1 + \mathfrak{N})$, the Lie algebra of $F \cap M$ is fully reducible by Lemma 4.4. Hence $F \cap M$ is fully reducible. By the opening remark of our proof, $F \cap M$ is a maximal fully reducible sub-

algebra of the algebraic group F. If $\dim \mathfrak{F}$ were less than $\mathfrak{G}$, we could apply the induction hypothesis to F and carry M_1 into $F \cap M$ by an inner automorphism from N and we would be led to the contradiction that $F \cap M$ is a maximal fully reducible subgroup of G. Consequently $\mathfrak{F} = \mathfrak{G}$, that is $\mathfrak{M}_1 + \mathfrak{N} = \mathfrak{M} + \mathfrak{N}$. From this it follows readily that the fully reducible subalgebra $\mathfrak{M}_1$ is maximal fully reducible in $\mathfrak{G}$.

By Theorem 4.1, $\mathfrak{M}_1$ and $\mathfrak{M}$ are conjugate under an inner automorphism from $\mathfrak{N}$. Hence no generality is lost in asuming $\mathfrak{M} = \mathfrak{M}_1$. Under this additional hypothesis, M_1 and M are each contained in M^A, the normalizer of $\mathfrak{M}$ in G. Furthermore the set of nilpotent elements in the radical of M^A is in $\mathfrak{N}$. Hence no generality is lost in assuming $G = M^A$. Repeating the 3-case argument in the proof of Theorem 6.1, we reduce to the case that $\mathfrak{G} = \mathfrak{N}$ is an abelian Lie algebra of nilpotent endomorphisms. If M_1 is a maximal fully reducible subgroup of G, $\mathfrak{M}_1$ consists of, in this case, fully reducible nilpotent elements and hence is 0. Thus M_1 is a finite group.[6] To prove that $u^{-1}M_1u \subset M$ for some u in N is equivalent to showing $M_1u \subset uM$; that is the group M_1 operating on the space of cosets G/M by left translation admits some fixed point uM. Since $G = M \cdot N$ (semi-direct), G/M can be identified with the space N. Under this identification, the operation of G on G/M by left translation is equivalent to the operation of G on the linear space $\mathfrak{N}$ by affine transformations in view of the identities

(1) $\quad mnM = (mnm^{-1}) \cdot M \qquad$ for m in M

(2) $\quad n_1(nM) = (n_1n) \cdot M \qquad$ for n_1 in N

(3) $\quad m \exp Xm^{-1} = \exp mXm^{-1} \qquad$ for X in N

(4) $\quad \exp(X+Y) = \exp X \cdot \exp Y \qquad$ for X, Y in $\mathfrak{N}$

together with the fact that $X \to \exp X$ is a one-to-one mapping of $\mathfrak{N}$ onto N. Now as is well-known, any finite group of affine transformations of a linear space (over a field of characteristic not dividing the order of the group) has a fixed point—namely, the centroid of any orbit. Hence M_1 has a fixed point on G/M and M_1 is conjugate to M under an inner automorphism from N. Proof of the theorem is now complete.

COROLLARY. *A pre-fully reducible group of similarities of an algebraic group keeps invariant a maximal fully reducible subgroup.*

---

[6] At this point the theorem follows from the fact that the one dimensional cohomology group of a finite group with coefficients in an indefinitely divisible abelian group is zero. Indeed this is in essence the concluding argument of our proof.

Proof. Let G be the algebraic group. The pre-fully reducible group of similarities of G are the automorphisms $g \to fgf^{-1}$ where f varies over a fully reducible group F. Let G_1 be the normalizer of G, let M be a maximal fully reducible subgroup of G and let M_1 be a maximal fully reducible subgroup of G_1 which includes M. Since $F \subset G_1$, there is an element x in G_1 with $F \subset xM_1x^{-1}$ by Theorem 7.1. Hence for any $f \varepsilon F$ and $p \varepsilon xMx^{-1}$, $fpf^{-1} \varepsilon (xM_1x^{-1}) \cap G$. But xM_1x^{-1} being a fully reducible subgroup of G_1, its intersection with G is a fully reducible subgroup of G (cf. 4.3) and therefore coincides with xMx^{-1}. Thus F keeps invariant the maximal fully reducible subgroup xMx^{-1}.

Section 8. Relation to Wedderburn Decomposition.

THEOREM 8.1. *Let G be an algebraic group and let $\mathfrak{E}$ denote its enveloping associative algebra. Then there is a Wedderburn decomposition $\mathfrak{S} + \mathfrak{X}$ for $\mathfrak{E}$ with $\mathfrak{S}$ semi-simple and $\mathfrak{X}$ the radical such that:*

a) *$\mathfrak{S} \cap G$ is a maximal fully reducible subgroup of G,*

b) *$\mathfrak{X} \cap \mathfrak{G}$ is the ideal of nilpotent elements in the radical of $\mathfrak{G}$, the Lie algebra of G,*

c) *$(I + \mathfrak{X}) \cap G$ is the subgroup of unipotent elements in the radical of G (I denotes the identity element).*

Proof. Let $G = M \cdot N$ be a decomposition of G into the semi-direct product of a maximal fully reducible subgroup M and the subgroup N of the unipotent elements in the radical.

For any set of endomorphisms S we denote by $\mathfrak{E}(S)$ the enveloping associative algebra—i. e. polynomial combinations without constant term in the elements of S. We denote by $\mathfrak{L}(S)$ the linear span of S. If S is a group then $\mathfrak{E}(S)$ consists only of linear combinations of elements of S. Hence $M \cdot N = N \cdot M$ implies that

$$\mathfrak{L}(\mathfrak{E}(M) \cdot \mathfrak{E}(N)) = \mathfrak{L}(\mathfrak{E}(N) \cdot \mathfrak{E}(M)) = \mathfrak{E}(M \cdot N) = \mathfrak{E}(G).$$

We denote by $\log u$ for u a unipotent endomorphism the finite sum

$$\sum_{i=1}^{\infty} (-1)^{i+1} (u-1)^i / i.$$

Since $u = \exp \log u$, we have

$$\mathfrak{E}(N) = \mathfrak{E}(\log N + I) = \mathfrak{E}(\log N) + KI = \mathfrak{E}(\mathfrak{N}) + KI,$$

where $\mathfrak{N} = \log N$ is the Lie algebra of N and K is the ground field. As a result

$$\mathfrak{E}(G) = \mathfrak{L}(\mathfrak{E}(M) \cdot (\mathfrak{E}(\mathfrak{N}) + KI)) = \mathfrak{E}(M) + \mathfrak{L}(\mathfrak{E}(M) \cdot \mathfrak{E}(N)).$$

Now $\mathfrak{N}$ being an ideal in $\mathfrak{G}$, we have $mnm^{-1} \varepsilon \mathfrak{N}$ whenever $m \varepsilon M$ and $n \varepsilon \mathfrak{N}$ and as a result

$$M \cdot \mathfrak{N} = \mathfrak{N} \cdot M, \mathfrak{L}(\mathfrak{E}(M) \cdot \mathfrak{E}(\mathfrak{N})) = \mathfrak{L}(\mathfrak{E}(\mathfrak{N}) \cdot \mathfrak{E}(M)),$$

and from this it follows that $\mathfrak{L}(\mathfrak{E}(M) \cdot \mathfrak{E}(\mathfrak{N}))$ is a two-sided ideal in the associative algebra $\mathfrak{E}(G)$. By Lie's theorem on the simultaneous triangularizing of solvable linear Lie algebras (or by Engel's theorem) $\mathfrak{N}^k = 0$ where k is the dimension of the linear space on which G operates. Hence $\mathfrak{E}(\mathfrak{N})^k = (0)$ and $\mathfrak{L}(\mathfrak{E}(M) \cdot \mathfrak{E}(\mathfrak{N}))^k = \mathfrak{L}(\mathfrak{E}(M) \cdot \mathfrak{E}(\mathfrak{N})^k) = 0$. Set $S = \mathfrak{E}(M)$ and $\mathfrak{T} = \mathfrak{L}(\mathfrak{E}(M) \cdot \mathfrak{E}(\mathfrak{N}))$. $\mathfrak{T}$ is in the radical of $\mathfrak{E}(\mathfrak{G})$. Moreover M being fully reducible, $\mathfrak{E}(M)$ is fully reducible and is therefore a semi-simple associative algebra. Thus $\mathfrak{S} \cap \mathfrak{T} = (0)$, and $\mathfrak{E}(\mathfrak{G}) = \mathfrak{S} + \mathfrak{T}$ (semi-direct). That is, $\mathfrak{S} + \mathfrak{T}$ is a Wedderburn decomposition of $\mathfrak{E}(G)$. Thus assertion a) has been proved.

To prove b), we observe that $\mathfrak{T} \cap \mathfrak{G}$ is invariant under inner automorphisms $y \to gyg^{-1}$ with g in G and it is therefore invariant under the infinitesimal transformations $y \to [g, y]$ with g in the Lie algebra $\mathfrak{G}$. Thus $\mathfrak{T} \cap \mathfrak{G}$ is an ideal of the Lie algebra $\mathfrak{G}$. Since it consists of nilpotent elements, it is a nilpotent ideal and therefore contained in the radical of $\mathfrak{G}$ and hence $\mathfrak{T} \cap \mathfrak{G} \subset \mathfrak{N}$. But by definition of $\mathfrak{T}$, $\mathfrak{N} \subset \mathfrak{T} \cap \mathfrak{G}$. Therefore $\mathfrak{T} \cap \mathfrak{G} = \mathfrak{N}$.

To prove c), we recall that $u = \exp \log u$ for u in N. Hence $N \subset I + \mathfrak{E}(\mathfrak{N})$ and $(I + \mathfrak{T}) \cap G \supset N$. Conversely, $u \varepsilon (I + \mathfrak{T}) \cap G$ implies $\log u \varepsilon \mathfrak{T}$ and also $\log u \varepsilon \mathfrak{G}$, the latter since $\{\exp t \log u \mid t \varepsilon K\}$ is the smallest algebraic group containing u. Hence $\log u \varepsilon \mathfrak{T} \cap \mathfrak{G} = \mathfrak{N}$ and $u \varepsilon N$. Proof of the theorem is now complete. Analogously, one can prove

THEOREM. *Let $\mathfrak{G}$ be an algebraic Lie algebra and $\mathfrak{E}$ its associative enveloping algebra. Then there exists a Wedderburn decomposition $\mathfrak{S} + \mathfrak{T}$ for $\mathfrak{E}$ such that $\mathfrak{S} \cap \mathfrak{G}$ is a maximal fully reducible subalgebra of $\mathfrak{G}$ and $\mathfrak{T} \cap \mathfrak{G}$ is the ideal of nilpotent elements in the radical of $\mathfrak{G}$.*

THE JOHNS HOPKINS UNIVERSITY
AND
INSTITUTO DE MATEMATICA PURA E APLIOADA,
RIO DE JANEIRO, BRAZIL.

REFERENCES.

[1] A. Borel and G. D. Mostow, "On semi-simple automorphisms of Lie algebras," *Annals of Mathematics*, vol. 61 (1955), pp. 389-405.

[2] C. Chevalley, *Theorie des Groupes de Lie*, vol. 2, Hermann et Cie, Paris (1951).

[3] ———, *Theorie des Groupes de Lie*, vol. 3, Hermann et Cie, Paris (1955).

[4] ———, "Algebraic Lie algebras," *Annals of Mathematics*, vol. 48 (1947), pp. 91-100.

[5] Harish-Chandra, *Proceedings of the American Mathematical Society*, vol. 1 (1950), pp. 14-17.

[6] N. Jacobson, *Annals of Mathematics*, vol. 36 (1935), pp. 875-881; also *Proceedings of the American Mathematical Society*, vol. 2 (1951), pp. 105-113.

[7] A. Malcev, "Solvable Lie algebras," *American Mathematical Society Translation* (No. 27) from *Izvestima Akademii Nauk SSSR*, vol. 9 (1945), pp. 329-352.

[8] G. D. Mostow, "Factor spaces of solvable groups," *Annals of Mathematics*, vol. 60 (1954), pp. 1-27.

[9] H. Weyl, *The Classical Groups*, Princeton University Press (1946).

[10] J. H. C. Whitehead, "On the decomposition of an infinitesimal group," *Proceedings of Cambridge Philosophical Society*, vol. 32 (1936), pp. 229-237.

[11] ———, "On certain equations on a group," *Quarterly Journal of Mathematics* (Oxford ser.), vol. 8 (1937), pp. 220-237.

CORRECTIONS TO "PERIODIC MAPPINGS ON A BANACH ALGEBRA." *

By Bertram Yood.

Dr. H. Kamel has kindly pointed out that the proof of Lemma 4.1 part (2) of this paper (this Journal, vol. 77, pp. 17-28) is incorrect. We have been unable to supply a correct proof. Although this lemma is used in the subsequent arguments we show that all the conclusions of the paper mentioned in the introduction hold. Only some subsidiary results require modification.

Consider Theorem 4.6. The proof for $n=1$ on p. 25 is valid since T has period two. For typographical convenience set $H_{2^j}(K_{2^j}) = H^j(K^j)$. We show by induction that $S \subset K^i$, $i = 0, 1, \cdots, n$ (whence $S = (0)$). This holds for $i=0$ by hypothesis and assume it for $0 \leqq i < n$. Set $V = T^{2^{i+1}}$. Let $\mathfrak{J}$ be any two-sided ideal of B where $\mathfrak{J} \subset K^i$. Since V is an automorphism, H^{i+1} is an algebra with $\mathfrak{J} \cap H^{i+1}$ as a two-sided ideal. By arguments of p. 25 given for J we have $\mathfrak{J} \cap H^{i+1} = (0)$. If $i+1 = n$ then we already have $S \cap H^n = (0)$ and $S = (0)$. Suppose that $i+1 < n$. Then

$$V(K^i) = (T^{2^i})^2(K^i) = K^i$$

by Lemma 2.1. Let r be the period of V. The algebraic sum

$$\mathfrak{J} = S + V(S) + \cdots + V^{r-1}(S)$$

is a two-sided ideal of B (since S is a two-sided ideal) and $\mathfrak{J} \subset K^i$ whence $\mathfrak{J} \cap H^{i+1} = (0)$. Let $y \in S$, $y = u + v$ in the decomposition

$$K^i = H^{i+1} \cap K^i \oplus K^{i+1}$$

of Lemma 2.2. Set $W = I + V + V^2 + \cdots + V^{r-1}$. Then $W(y) \in \mathfrak{J}$, $W(v) = 0$ and $W(u) = ru$. Thus $u = 0$ and $y = v \in K^{i+1}$. This completes the induction.

This gives the correctness of all the results subsequent to Theorem 4.6 with the exception of Theorem 4.10. For that result we argue as follows. $\| x \|$ and $\| x \|_1 = \| T(x) \|$ are two complete norms on S. Let x_i be in separating ideal S_1 for these norms on S, and let $\| v_i - x_i \| \to 0$, $\| v_i \|_1 \to 0$ where $v_i \in S \subset K_1$, $i = 1, 2$. Arguing as in Lemma 4.9 we see that $x_1 x_2 = 0$. Thus

* Received August 30, 1955.

S_1 is a zero algebra. Now S is semi-simple being a two-sided ideal in B. Thus $S_1 = (0)$. By Theorem 2.4, S is a zero algebra whence $S = (0)$.

LEMMA. *Let T be a periodic automorphism or anti-automorphism on a Banach algebra B with separating ideal S. If T^2 is continuous then $T(S) = S$. If T^2 is continuous on $\bar{K}_1$, $u + v \varepsilon S$ where $u \varepsilon H_1$, $v \varepsilon K_1$, then $u \varepsilon S$ and $v \varepsilon S$.*

Suppose T^2 continuous. Let $y = u + v \varepsilon S$, $u_k + v_k \to y$, $u_k + T(v_k) \to 0$ where each u_k and u (v_k and v) lies in $H_1(K_1)$. Then $u_k + T^2 v_k \to u + T^2(v)$. Also $z_k = u - u_k + T(v - v_k) \to T(y)$ while $T(z_k) \to 0$. Thus $T(S) \subset S$ and by periodicity $T(S) = S$.

Suppose T^2 continuous on $\bar{K}_1$. Suppose first that the period n of T is even. Let $y \varepsilon S$ with the above notation. Then $w_k = v_k - v - T(v_k) \to u$ with $w_k \varepsilon K_1$ by Lemma 2.1. Then $u \varepsilon K_1' \cap H_1 \subset S$ by Lemma 4.2(b). Hence also $v \varepsilon S$. If n is odd then $T = T^{n+1}$ is continuous on $\bar{K}_1$. Now $v_k - T(v_k) \to u + v$. Operating by T on this $n - 1$ times and summing we obtain $nu = 0$. Thus $v \varepsilon S$.

The additional hypothesis $T(S) = S$ (or T^2 continuous) yields 4.1(3), (4), 4.2(a), 4.3 and 4.4(a). The additional hypothesis that T^2 is continuous on $\bar{K}_1$ yields 4.1(3), (4). For 4.4(b) assume $T(S) = S$ in case $n = 4$.

UNIVERSITY OF OREGON.

CORRECTIONS TO THE PAPER "ENGEL RINGS AND A RESULT OF HERSTEIN AND KAPLANSKY." *

By M. P. DRAZIN.

In the above-named paper (this JOURNAL, vol. 77 (1955), pp. 895-913), Theorem 6.4 should have been stated only for rings of zero characteristic: the argument for the case of prime characteristic breaks down in the last formula line on p. 911. This involves jettisoning Theorem 6.5 (the implied "proof" of which depended on applying Theorem 6.4 to homomorphs which cannot be guaranteed to have zero characteristic, even when the given ring has).

However, the writer has no evidence that Theorem 6.4 as stated or Theorem 6.5 is actually false. And in any case, since the valid part of the proof of Theorem 6.4 establishes the weaker conclusion $[x^m, y^{np^r}] = 0$ *without* characteristic hypothesis, the remarks following Theorem 6.4 still hold good, provided that we modify the final parenthetical clause so as to read "which would at any rate imply that every weak K-ring R with k, m, n satisfying (a) or (b) has R/J commutative."

TRINITY COLLEGE, CAMBRIDGE.

* Received December 6, 1955.

SUBALGEBRAS OF THE ALGEBRA OF ALL COMPLEX-VALUED CONTINUOUS FUNCTIONS ON THE CIRCLE.*

By JOHN WERMER.[1]

1. Introduction. Let $\boldsymbol{R}$ be the algebra of all real valued continuous functions on the circle and let $\boldsymbol{C}$ be the algebra of all complex-valued continuous functions on the circle.

The subalgebras we are concerned with are assumed to contain the constant 1.

A fundamental theorem of Stone yields that any uniformly closed subalgebra $\boldsymbol{R}'$ of $\boldsymbol{R}$ which separates points (i. e which is such that if $\lambda_1 \neq \lambda_2$ then there exists an f in $\boldsymbol{R}'$ with $f(\lambda_1) \neq f(\lambda_2)$) coincides with $\boldsymbol{R}$.

For the algebra $\boldsymbol{C}$ the situation is quite different. There exists a large class of proper subalgebras of $\boldsymbol{C}$ which separate points. The problem of classifying these subalgebras leads to the following question: What are the *maximal* subalgebras of $\boldsymbol{C}$?

A closed proper subalgebra $\boldsymbol{M}$ *of* $\boldsymbol{C}$ separating points is called *maximal* if there exists no closed subalgebra $\boldsymbol{M}'$ with $\boldsymbol{M} \subseteq \boldsymbol{M}'$ and $\boldsymbol{M}' \neq \boldsymbol{M}$, $\boldsymbol{M}' \neq \boldsymbol{C}$.

In [1] and [2] the author has given examples of certain maximal subalgebras. Here we shall exhibit a large class of maximal subalgebras, associated with Riemann surfaces.

Let $\mathfrak{F}$ be a Riemann surface,[2] $\mathfrak{M}$ a region on $\mathfrak{F}$ bounded by a simple closed analytic curve γ, such that $\mathfrak{M} + \gamma$ is compact. $\mathfrak{M}$ then has finite genus p. Since γ is topologically a circle, we may regard $\boldsymbol{C}$ as the space of continuous complex-valued functions on γ. We shall use $\| f \|$ to denote $\max_{\lambda \varepsilon \gamma} | f(\lambda) |$.

Definition 1. $\mathfrak{A}$ is the subalgebra of $\boldsymbol{C}$ consisting of all f which may be continued into $\mathfrak{M}$ to be analytic on $\mathfrak{M}$ and continuous on $\mathfrak{M} + \gamma$.

By the maximum principle for $\mathfrak{M}$, $\mathfrak{A}$ is a closed subalgebra of $\boldsymbol{C}$. Also

* Received September 10, 1954.

[1] I am indebted to Professor M. Heins for a number of valuable suggestions which I have used in some of the proofs.

[2] For definitions and basic facts of the theory of Riemann surfaces we refer the reader to R. Nevanlinna, "Uniformisierung," Springer Verlag, 1953.

$\mathfrak{A}$ separates points on γ, since, given p, q on γ, $p \neq q$, we can find some f in $\mathfrak{A}$ with $f(p) = 0$, $f(q) \neq 0$.

The main object of this paper is to prove:

THEOREM 2. *$\mathfrak{A}$ is a maximal subalgebra of $\boldsymbol{C}$.*

When $\mathfrak{F}$ is the plane and γ is the unit circle $|\lambda| = 1$, $\mathfrak{A}$ becomes the algebra generated by the functions 1 and λ. This is the case discussed in [1].

In Section 2 we prove Theorem 1 in which we give the form of the general linear functional on $\boldsymbol{C}$ which annihilates $\mathfrak{A}$. In Section 3 we use Theorem 1 to prove Theorem 2. In Section 4 we find when two of our maximal subalgebras are isomorphic.

2. Fix ζ in $\mathfrak{M}$; let G_ζ denote the Green's function for $\mathfrak{M}$ singular at ζ. Then G_ζ is harmonic in $\mathfrak{M}$ except at ζ; G_ζ vanishes on γ; for some fixed local parameter z at ζ, $G_\zeta(z) + \log|z - \zeta|$ is regular neighborhood of ζ.

Let H_ζ be the (multiple-valued) conjugate function of G_ζ. Since γ is an analytic curve, $G_\zeta + iH_\zeta$ is analytic everywhere on γ.

Set $W_\zeta(z) = -d\{G_\zeta(z) + iH_\zeta(z)\}/dz$. Then W_ζ is a "covariant" on $\mathfrak{M}$, as defined in [3], p. 102. W_ζ is analytic on $\mathfrak{M}$ except for a simple pole at ζ, with residue 1. On γ we denote by ω_ζ the measure $\frac{1}{2\pi i} W_\zeta(\lambda)\, d\lambda$. Then ω_ζ is the harmonic measure for $\mathfrak{M}$ evaluated at ζ. In particular, for any set E on γ, $\omega_\zeta(E)$ is real and non-negative.

Fix some ζ_0 in $\mathfrak{M}$. From now on we shall omit the subscript ζ_0 when writing G_{ζ_0}, H_{ζ_0}, W_{ζ_0} or ω_{ζ_0}. We note that W has no zero on γ.

LEMMA 1. *W has $2p$ zeros in $\mathfrak{M}$, where each zero is counted with its multiplicity.*

Proof. We choose an analytic parametrization of $\gamma: \lambda = \lambda(t)$, $0 \leqq t \leqq 1$. In any coordinate neighborhood U of a point of γ with local parameter z, we define $z(t) = z(\lambda(t))$. We can then consider $W(z(t)) \cdot z'(t)$ in $U \cap \gamma$, where the prime indicates differentiation with respect to t. Direct computation shows that this expression is independent of the choice of local parameter. Then

$$\Delta = \int_0^1 d/dt\{\log W(z(t))z'(t)\}dt$$

is well-defined. By a formula given in [3], p. 133, we have

$$\Delta = 2\pi i(B - A - N),$$

where B is the number of zeros of W in $\mathfrak{M}$, A the number of poles, and N the Euler characteristic of $\mathfrak{M}$. On the other hand

$$\Delta = \int_0^1 d/dt \log d/dt - \{G(z(t)) + iH(z(t))\} dt$$

and this equals the variation of $\log d - \{G(z(t)) + iH(z(t))\}/dt$ over $0 \leqq t \leqq 1$. The properties of G and H on γ then yield directly that $\Delta = 0$. Thus $B = A + N$. But $A = 1$ and $N = 2p - 1$. Hence $B = 2p$, as asserted.

Let $\alpha_1, \cdots, \alpha_{2p}$ be a homology basis of closed curves on $\mathfrak{M}$.

Definition 2. Let $u(\lambda)$ be a real continuous function on γ and let $U(\zeta)$ be the harmonic function on $\mathfrak{M}$ with $U \equiv u$ on γ. Then $\Phi_\nu(u)$, $\nu = 1, 2, \cdots, 2p$, denotes the period of the conjugate function of U corresponding to α_ν.

Definition 3. For $\mu = 1, \cdots, 2p$, $\psi_\mu(\zeta)$ is a harmonic function on $\mathfrak{M}$, continuous on $\mathfrak{M} + \gamma$ and twice differentiable on γ, with $\Phi_\nu(\psi_\mu) = \delta_\mu{}^\nu$.

LEMMA 2. *There exist functions K_i, $i = 1, \cdots, 2p$, meromorphic on $\mathfrak{M}$ and continuous and real valued on γ such that for $i = 1, \cdots, 2p$, $\Phi_i = K_i d\omega$ as functionals, i. e. $\Phi_i(u) = \int_\gamma u(\lambda) K_i(\lambda) d\omega(\lambda)$, all u, and such that $W(\zeta)K_i(\zeta)$ is analytic on $\mathfrak{M}$ for each i.*

Proof. Let W have the zeros $z_1, \cdots, z_k$ in $\mathfrak{M}$ of orders $\nu_1, \cdots, \nu_k$. By Lemma 1, $\sum_{i=1}^{i=k} \nu_i = 2p$. At z_i we use local polar coordinates (r_i, θ_i) for $i = 1, \cdots, k$.

For each i, $i = 1, \cdots, k$ a well-known construction yields us functions $u_j{}^i$, $j = 1, \cdots, \nu_i$ and $v_j{}^i$, $j = 1, \cdots, \nu_i$ such that:

(i) $u_j{}^i$, $v_j{}^i$ are harmonic on $\mathfrak{M}$ except at z_i.

(ii) $u_j{}^i$ has at z_i a pole with principal part $(r_i)^{-j} \cos j\theta_i$ and $v_j{}^i$ has at z_i a pole with principal part $(r_i)^{-j} \sin j\theta_i{}^j$.

(iii) $u_j{}^i$ and $v_j{}^i$ vanish identically on γ.

The $4p$ functions $u_j{}^i$, $v_j{}^i$ together form a linearly independent set. For suppose $\sum_{i,j} c_j{}^i u_j{}^i + d_j{}^i v_j{}^i = 0$, $c_j{}^i$, $d_j{}^i$ being constants. Because of the poles of the $u_j{}^i$, $v_j{}^i$ this implies $c_j{}^i u_j{}^i + d_j{}^i v_j{}^i = 0$, whence $c_j{}^i \cos j\theta_i + d_j{}^i \sin j\theta_i = 0$. Since θ_i is arbitrary, we conclude $c_j{}^i = d_j{}^i = 0$, all i, j. Thus linear independence is established. We now seek constants $a_j{}^i$, $b_j{}^i$, $j = 1, \cdots, \nu_i$, $i = 1, \cdots, 2p$,

such that the function $\sum_{i,j} a_j{}^i u_j{}^i + b_j{}^i v_j{}^i$ have a single-valued conjugate function. This gives $2p$ conditions on $4p$ unknowns, and so we obtain at least $2p$ linearly independent $4p$-tuples satisfying the conditions. We thus get the linearly independent functions

$$q^\nu(\zeta) = \sum_{i,j} a_{j\nu}{}^i u_j{}^i(\zeta) + b_{j\nu}{}^i v_j{}^i(\zeta), \qquad \nu = 1, \cdots, 2p.$$

where each q^ν has a single-valued conjugate function. Then $q^\nu(\zeta)$ is harmonic on $\mathfrak{M}$ except for possible poles at the points z_i of orders $\leqq \nu_i$ and $q^\nu(\lambda) = 0$ on γ for all ν.

Let $r^\nu(\zeta)$ be the conjugate function of q^ν with $r^\nu(\zeta_0) = 0$. Then $h_\nu(\zeta) = i(q^\nu(\zeta) + ir^\nu(\zeta))$ is, for $\nu = 1, \cdots, 2p$, a meromorphic function on $\mathfrak{M}$ such that $h_\nu(\lambda)$ is real and continuous on γ for each ν. Suppose $\sum_{\nu=1}^{2p} c_\nu h_\nu = 0$, where $c_\nu = a_\nu + ib_\nu$, (a_ν, b_ν real). Since $h_\nu(\lambda)$ is real for λ and γ, this gives $\sum_{\nu=1}^{2p} a_\nu h_\nu(\lambda) = \sum_{\nu=1}^{2p} b_\nu h_\nu(\lambda) = 0$ for λ in γ. But $\sum_{\nu=1}^{2p} a_\nu h_\nu(\zeta)$ is meromorphic and so $\sum_{\nu=1}^{2p} a_\nu h_\nu(\zeta) = 0$, $\zeta \varepsilon \mathfrak{M}$. Similarly $\sum_{\nu=1}^{2p} b_\nu h_\nu(\zeta) = 0$ on $\mathfrak{M}$. Hence $\sum_{\nu=1}^{2p} a_\nu q^\nu(\zeta) = \sum_{\nu=1}^{2p} b_\nu q^\nu(\zeta) = 0$ on $\mathfrak{M}$. But the q^ν are linearly independent by construction, whence $a_\nu = b_\nu = 0$, all ν. Hence $c_\nu = 0$, all ν. Hence the h_ν are linearly independent.

Consider now the covariant $h_\nu(\zeta) W(\zeta)$. This has no poles except possibly at ζ_0 since the zeros of W cancel the poles of h_ν. Also $h_\nu(\zeta_0) = iq^\nu(\zeta_0)$ and so $\frac{1}{2\pi i} \int_\gamma h_\nu(\lambda) W(\lambda) d\lambda = iq^\nu(\zeta_0)$ by the residue theorem. Now $h_\nu(\lambda)$ is real on γ, and $\frac{1}{2\pi i} W(\lambda) d\lambda$ is a real-valued measure on γ. Hence the left hand side is real. Hence $q^\nu(\zeta_0) = 0$ and so $h_\nu(\zeta_0) = 0$. It follows that $h_\nu(\zeta) W(\zeta)$ is regular at ζ_0 and so everywhere on $\mathfrak{M}$.

Let now N be the space of all real continuous functions u on γ with $\Phi_\nu(u) = 0$ for $\nu = 1, \cdots, 2p$. Given u in N choose v twice differentiable on γ with $\| u - v \| < \epsilon$. Then there is a constant K so that

$$| \Phi_\nu(v) | = | \Phi_\nu(v - u) | \leqq K\epsilon, \qquad \nu = 1, \cdots, 2p.$$

Set $w(\lambda) = v(\lambda) - \sum_{\nu=1}^{2p} \Phi_\nu(v) \psi_\nu(\lambda)$. Then w is differentiable on γ, $\Phi_j(w) = 0$, $j = 1, \cdots, 2p$ and

$$\| u - w \| \leqq \| u - v \| + \sum_{\nu=1}^{2p} | \Phi_\nu(v) | \, \| \psi_\nu \| < \epsilon + K'\epsilon = K''\epsilon, K''$$

independent of ϵ.

Let $w(\zeta)$ be the harmonic function with boundary value $w(\lambda)$. Since $\Phi_j(w)=0$, $j=1,\cdots,2p$, w has a single-valued conjugate w_1 and since w is twice differentiable on γ, w_1 is continuous in $\mathfrak{M}+\gamma$. By the preceding $h_i(\zeta)W(\zeta)$ is analytic on $\mathfrak{M}$ for each i. The residue theorem yields:

$$0=\int_\gamma (w(\lambda)+iw_1(\lambda))h_i(\lambda)W(\lambda)d\lambda,$$

all i. Since $h_i(\lambda)$ is real, we get

$$0=\int_\gamma w(\lambda)h_i(\lambda)W(\lambda)d\lambda=\int_\gamma w(\lambda)h_i(\lambda)dw(\lambda)$$

for all i. Now

$$\begin{aligned}\Big|\int_\gamma u(\lambda)h_i(\lambda)dw(\lambda)\Big| &= \Big|\int_\gamma (u(\lambda)-w(\lambda))h_i(\lambda)dw(\lambda)\Big|\\ &\leqq K''\epsilon\int_\gamma |h_i(\lambda)|\,dw(\lambda).\end{aligned}$$

Since ϵ is arbitrary, we get

$$\int_\gamma u(\lambda)h_i(\lambda)dw(\lambda)=0,$$

all i. Thus the functional $h_i dw$ annihilates N. Hence, by elementary vector-space reasoning, there exist constants $b_\nu{}^i$, $i=1,\cdots,2p$, $\nu=1,\cdots,2p$, with

$$h_i(\lambda)dw(\lambda)=\sum_{\nu=1}^{2p} b_\nu\Phi_\nu, \qquad i=1,\cdots,2p.$$

Since the h_i are linearly independent, we can solve this system of equations to get

$$\Phi_\nu=\sum_{i=1}^{2p} c_\nu{}^i h_i(\lambda)dw(\lambda)=K_\nu(\lambda)dw(\lambda), \qquad \nu=1,\cdots,2p.$$

The properties of the h_i established above yield that the K_i satisfy the assertions of the Lemma.

LEMMA 3. *Let μ be any complex-valued Borel measure on γ such that $\int_\gamma f(\lambda)d\mu(\lambda)=0$ whenever $f\varepsilon\mathfrak{A}$. Then for closed sets E on γ, $\omega(E)=0$ implies $\mu(E)=0$.*

The analogous assertion was proved for the unit circle by F. and M. Riesz in [4]. A slight modification of their argument yields the following proof.

Proof of Lemma 3. Since E is closed, the complement of E on γ is the union of countably many disjoint arcs γ_n. Since

$$\int_\gamma d\omega(\lambda) < \infty, \qquad \sum_{n=1}^{\infty} \int_{\gamma_n} d\omega(\lambda) < \infty.$$

Hence we can find a sequence of positive numbers d_n with d_n increasing to infinity with n, such that $\sum_{n=1}^{\infty} (\int_{\gamma_n} d\omega(\lambda)) \cdot d_n < \infty$. For $n = 1, 2, \cdots$ we define a positive real twice differentiable function g_n on γ_n such that $I_n = \int_{\gamma_n} g_n(\lambda) d\omega(\lambda) < \infty$ and $g_n(\lambda)$ increases on γ_n to ∞ as λ approaches the endpoints of γ_n. Choose positive constants c_n with $\sum_{n=1}^{\infty} c_n I_n < \infty$.

Set $P(\lambda) = c_n g_n(\lambda) + d_n$ for $\lambda \varepsilon \gamma_n$, $n = 1, 2, \cdots$. Since $\omega(E) = 0$, P is defined almost everywhere on γ with respect to ω and so with respect to ω_ζ for every ζ in $\mathfrak{M}$. Let now $P(\zeta) = \int_\gamma P(\lambda) d\omega_\zeta(\lambda)$, $\zeta \varepsilon \mathfrak{M}$. Note that

$$\int_\gamma P(\lambda) d\omega(\lambda) = \sum_{n=1}^{\infty} (c_n I_n + d_n \int_{\gamma_n} d\omega(\lambda)) < \infty.$$

From the way $P(\lambda)$ was constructed, we see that $P(\lambda)$ is continuous and finite at each point $\lambda \not\varepsilon E$ and $P(\lambda)$ becomes continuously $+\infty$ at each point of E.

By elementary properties of the harmonic measures ω_ζ, we get then that $P(\zeta)$ is harmonic in $\mathfrak{M}$ and has $P(\lambda)$ as continuous boundary function on γ, if we make the obvious definition of continuous approach to ∞ for $P(\zeta)$ as $\zeta \to \lambda$, λ in E. Let $k_1, \cdots, k_{2p}$ be the periods of the conjugate function of $P(\zeta)$. Choose a constant d so that $P_1 = d + P - \sum_{i=1}^{2n} k_i \psi_i > 0$ on γ, and hence on $\mathfrak{M}$ and let Q_1 be the (single-valued) conjugate function of P_1. Set now $k(\zeta) = (1 + P_1 + iQ_1)^{-1} \cdot (P_1 + iQ_1)(\zeta)$. Then $k(\zeta)$ is analytic in $\mathfrak{M}$. For $\lambda \varepsilon \gamma$ and $\lambda \not\varepsilon E$, $k(\lambda) = (1 + x + iy)^{-1} \cdot (x + iy)$ with $0 < x < \infty$ and hence $|k(\lambda)| < 1$. Let $\lambda \varepsilon E$. As $\zeta \to \lambda$ $|P_1(\zeta) + iQ_1(\zeta)| \to \infty$, whence $k(\zeta) \to 1$. Thus $k(\lambda) = 1$ on E. In particular $k(\lambda)$ is continuous in $\mathfrak{M} + \gamma$, and so $k \varepsilon \mathfrak{A}$. Hence $k^n \varepsilon \mathfrak{A}$ for $n = 1, 2, \cdots$. Then

$$0 = \int_\gamma k^n(\lambda) d\mu(\lambda) = \int_E d\mu(\lambda) + \int_{\gamma - E} k^n(\lambda) d\mu(\lambda).$$

Letting $n \to \infty$ and recalling that $|k(\lambda)| < 1$ if $\lambda \varepsilon \gamma - E$, we conclude that $0 = \int_E d\mu(\lambda) = \mu(E)$, as asserted.

Definition 4. We denote by $L^p(\gamma)$ the class of functions $F(\lambda)$ on γ measurable with respect to ω and with $\int_\gamma |F(\lambda)|^p d\omega(\lambda) < \infty$.

COROLLARY. *If μ is a Borel-measure on γ such that $\int_\gamma f(\lambda) d\mu(\lambda) = 0$ for all $f \varepsilon \mathfrak{A}$, then there exists $F(\lambda) \varepsilon L^1(\gamma)$ such that $d\mu(\lambda) = F(\lambda) d\omega(\lambda)$ as measures on γ.*

Proof. By Lemma 3, $\omega(E) = 0$ implies $\mu(E) = 0$ for any closed set E. It follows that this implication holds for each Borel set E.

We can write $\mu = \mu^+ - \mu^- + i\nu^+ - i\nu^-$ where μ^+, μ^-, ν^+, ν^- are real non-negative measures. Let now E be any Borel set with $\omega(E) = 0$. For every Borel subset E' of E, $\omega(E') = 0$ and so $\mu(E') = 0$. Hence $\mu^+(E) = \mu^-(E) = \nu^+(E) = \nu^-(E) = 0$. Thus μ^+, etc. are all absolutely continuous with respect to ω. It follows by the Radon-Nikodym theorem, that $d\mu^+(\lambda) = F_1(\lambda) d\omega(\lambda)$, where $F_1 \varepsilon L^1(\gamma)$. Similiarly $d\mu^-(\lambda) = F_2(\lambda) d\omega(\lambda)$, $F_2 \varepsilon L^1(\gamma)$ and so on. Adding these equations we get the assertion.

Let now γ' be a simple closed analytic curve in $\mathfrak{M}$ such that γ and γ' together bound an annular subregion $\mathfrak{M}'$ of $\mathfrak{M}$. We choose γ' so that all zeros of W lie outside $\mathfrak{M}' + \gamma'$. We can then map $\mathfrak{M}'$ conformally onto the annulus $r' < |z| < 1$ in the plane, by a mapping $z = \chi(\zeta)$, $\zeta \varepsilon \mathfrak{M}'$. Since γ, γ' are analytic curves, χ is analytic on the boundary curves γ and γ'. It follows that for a fixed K, and each Borel set E on γ,

$$\frac{1}{K}\,\omega(E) \leqq m(\chi(E)) \leqq K\omega(E),$$

where m denotes Lebesgue measure on $|z| = 1$.

Let $F(\zeta)$ be analytic on $\mathfrak{M}'$. Then $F^o(z) = F(\chi^{-1}(z))$ is analytic in the annulus $r' < |z| < 1$. We shall omit the symbol "o," since this omission introduces no ambiguity. Also for $g(\lambda)$ defined on γ, we write $g(e^{i\theta})$ for $g(\chi^{-1}(e^{i\theta}))$.

Definition 5. Let F be analytic in $\mathfrak{M}$. We say $F \varepsilon \mathfrak{H}'$, provided that (with the notations just given)

$$\int_0^{2\pi} |F(re^{i\theta})|\, d\theta = O(1) \quad as \ r \to 1.$$

LEMMA 4. *Let $F(\zeta) \varepsilon \mathfrak{H}'$. Then there exists a function $F^*(\lambda)$ defined on γ* a. e.—*$d\omega$ such that*

(a) $F^* \varepsilon L^1(\gamma)$.

(b) $F(\zeta)=\frac{1}{2\pi i}\int_{\gamma} F^*(\lambda)W_\zeta(\lambda)\,d\lambda$, *all* $\zeta\,\varepsilon\,\mathfrak{M}$.

(c) $\lim\limits_{\zeta\to\lambda} F(\zeta)=F^*(\lambda)$ a. e.—$d\omega$ *on* γ, *if* $\zeta\to\lambda$ *within some sector.*

(d) *Fix* $r_1>r'$. *Then for some constant* K *independent of* F *and* r *we have for* $r_1<r<1$:

$$\int_0^{2\pi}|F(re^{i\theta})|\,d\theta\leqq K\int_0^{2\pi}|F^*(e^{i\theta})|\,d\theta.$$

Proof. By hypothesis $\int_0^{2\pi}|F(re^{i\theta})|\,d\theta=O(1)$ as $r\to 1$, and $F(z)$ is analytic for $r'<|z|<1$. We may write $F=F_1+F_2$ where F_1 is analytic in $|z|<1$ and F_2 is analytic in $|z|>r'$. We hence get

$$\int_0^{2\pi}|F_1(re^{i\theta})|\,d\theta=O(1) \text{ as } r\to 1.$$

Classical results now give that $F_1^*(e^{i\theta})=\lim\limits_{z\to e^{i\theta}}F_1(z)$ exists for a. a. θ if the approach to the boundary lies within some sector, and that $F_1^*(e^{i\theta})$ is summable on $0\leqq\theta<2\pi$. Hence $\lim\limits_{z\to e^{i\theta}}F(z)$ exists a. e. We denote it by $F^*(e^{i\theta})$; for λ on γ we write $F^*(\lambda)$ instead of $F^*(\chi(\lambda))$. Since sets of ω-measure 0 on γ correspond to sets of Lebesgue measure 0 on $|z|=1$, we so get assertion (c).

Let now $r'<r<1$ and let γ_r be the curve in $\mathfrak{M}'$ which χ maps into the circle $|z|=r$. The residue theorem gives for $\zeta\,\varepsilon\,\mathfrak{M}$, ζ outside the region bounded by γ_r and γ:

$$F(\zeta)=\frac{1}{2\pi i}\int_{\gamma_r}F(\lambda)W_\zeta(\lambda)\,d\lambda=\frac{1}{2\pi}\int_0^{2\pi}F(re^{i\theta})W_\zeta(re^{i\theta})re^{i\theta}d\theta.$$

By a classical theorem, $\lim\limits_{r\to 1}\int_0^{2\pi}|F_1(re^{i\theta})-F_1^*(e^{i\theta})|\,d\theta=0$; also $W_\zeta(re^{i\theta})$ is continuous for $r'<r\leqq 1$. It follows that

$$F(\zeta)=\lim_{r\to 1}\frac{1}{2\pi}\int_0^{2\pi}F(re^{i\theta})W_\zeta(re^{i\theta})re^{i\theta}\,d\theta=\frac{1}{2\pi}\int_0^{2\pi}F^*(e^{i\theta})W_\zeta(e^{i\theta})e^{i\theta}\,d\theta.$$

Hence $F(\zeta)=\frac{1}{2\pi i}\int_\gamma F^*(\lambda)W_\zeta(\lambda)\,d\lambda$, $\zeta\,\varepsilon\,\mathfrak{M}$. Thus (b) is proved.

Now

$$\int_\gamma|F^*(\lambda)|\,d\omega(\lambda)=\frac{1}{2\pi}\int_0^{2\pi}|F^*(e^{i\theta})|\,W(e^{i\theta})e^{i\theta}\,d\theta<\infty,$$

since $\int_0^{2\pi} |F^*(e^{i\theta})|\, d\theta < \infty$ by construction of F^*, and W is continuous. Thus (a) holds. Finally, write $F(z) = F_1(z) + F_2(z)$, where $F_1(z)$ is analytic in $|z| < 1$ and $F_2(z) = -\frac{1}{2\pi i}\int_{|\tau|=r'} F(\tau)(\tau - z)^{-1} d\tau$; so that F_2 is analytic in $|z| > r'$. Let $\zeta = \chi^{-1}(\tau)$. Then $F(\tau) = \int_\gamma F^*(\lambda)(W(\lambda))^{-1} W_\zeta(\lambda) d\omega(\lambda)$ by (b), whence $|F(\tau)| \leqq M_\zeta \int_\gamma |F^*(\lambda)|\, d\omega(\lambda)$, where $M_\zeta = \max_{\lambda \varepsilon \gamma} |W^{-1}(\lambda) W_\zeta(\lambda)|$. Hence for $0 \leqq \theta < 2\pi$, $r' < r$,

$$|F_2(re^{i\theta})| \leqq 1/(r - r') \max_{|\tau|=r'} |F(\tau)| \leqq (M/2\pi) 1/(r - r') \int_\gamma |F^*(\lambda)|\, d\omega(\lambda),$$

where $M = \sup_{\zeta \varepsilon \gamma'} M_\zeta$. At last, $r' < r_1 \leqq r$,

$$\int_0^{2\pi} |F_2(re^{i\theta})|\, d\theta \leqq \int_\gamma |F^*(\lambda)|\, d\omega(\lambda) \cdot M_1 \leqq M_2 \cdot \int_0^{2\pi} |F^*(e^{i\theta})|\, d\theta,$$

where M_1 and M_2 are constants. On the other hand, since F_1 is analytic in $|z| < 1$,

$$\int_0^{2\pi} |F_1(re^{i\theta})|\, d\theta \leqq \int_0^{2\pi} |F_1^*(e^{i\theta})|\, d\theta \leqq \int_0^{2\pi} |F^*(e^{i\theta})|\, d\theta + \int_0^{2\pi} |F_2^*(e^{i\theta})|\, d\theta$$
$$\leqq (1 + M_2) \int_0^{2\pi} |F^*(e^{i\theta})|\, d\theta.$$

Hence

$$\int_0^{2\pi} |F(re^{i\theta})|\, d\theta = \int_0^{2\pi} |F_1(re^{i\theta}) + F_2(re^{i\theta})|\, d\theta$$
$$\leqq (1 + 2M_2) \int_0^{2\pi} |F^*(e^{i\theta})|\, d\theta.$$

This proves (d).

Definition 6. $\mathfrak{B}$ is the conjugate space of ***C***.

By the representation theorem of F. Riesz, $\mathfrak{B}$ may be identified with the space of all complex-valued Borel-measures on γ.

Definition 7. $\mathfrak{W}$ is the subspace of $\mathfrak{B}$ consisting of all measures μ of the form $d\mu(\lambda) = G^*(\lambda) d\omega(\lambda)$ where G^* is the boundary function of some G in $\mathfrak{H}'$ with $G(\zeta_0) = 0$.

LEMMA 5.[3] *$\mathfrak{W}$ is regularly closed as subspace of $\mathfrak{B}$.*

[3] The idea of using a lemma of this kind resulted from a conversation with Professor S. Kakutani.

Proof. By a theorem of Banach, [5], p. 124, it suffices to show that with each weakly convergent sequence of elements of $\mathfrak{W}$ the limit again is in $\mathfrak{W}$.

Let now $\mu_n \varepsilon \mathfrak{W}$, μ_n converge weakly to μ. By Definition 7, there exists $G_n \varepsilon \mathfrak{H}'$, $G_n(\zeta_0) = 0$ with $d\mu_n(\lambda) = G_n^*(\lambda) d\omega(\lambda)$. Then for each $f \varepsilon \mathfrak{A}$, $f \cdot G_n \varepsilon \mathfrak{H}'$, and so by Lemma 4,

$$0 = f(\zeta_0) G_n(\zeta_0) = \int_\gamma f(\lambda) G_n^*(\lambda) d\omega(\lambda) = \int_\gamma f(\lambda) d\mu_n(\lambda).$$

For $f \varepsilon \mathfrak{A}$, then, $0 = \int_\gamma f(\lambda) d\mu(\lambda)$. By the Corollary to Lemma 3, this implies that there exists $G_0 \varepsilon L^1(\gamma)$ with $G_0(\lambda) d\omega(\lambda) = d\mu(\lambda)$.

We shall show that G_n converges to a function G analytic on $\mathfrak{M}$ with $G \varepsilon \mathfrak{H}'$ and that $G^*(\lambda) = G_0(\lambda)$ a. e.—$d\omega$ on γ. From this it follows that $d\mu(\lambda) = G^*(\lambda) d\omega(\lambda)$ and so $\mu \varepsilon \mathfrak{W}$. Now

$$G_n(\zeta) = \frac{1}{2\pi i} \int_\gamma G_n^*(\lambda) W_\zeta(\lambda) d\lambda = \int_\gamma (W(\lambda))^{-1} W_\zeta(\lambda) d\mu_n(\lambda).$$

Hence $G(\zeta) = \lim_{n\to\infty} G_n(\zeta)$ exists for $\zeta \varepsilon \mathfrak{M}$. Also

$$| G_n(\zeta) | \leqq M_\zeta \int_\gamma | G_n^*(\lambda) | d\omega(\lambda),$$

where $M_\zeta = \max_{\lambda \varepsilon \gamma} | (W(\lambda))^{-1} W_\zeta(\lambda) |$. Now since the sequence μ_n converges weakly, the total variation of μ_n has a bound K valid for all n. Hence

$$\int_\gamma | G_n^*(\lambda) | d\omega(\lambda) < K. \tag{1}$$

Also M_ζ is bounded on each compact subset of $\mathfrak{M}$. Hence by Vitali's theorem, G is analytic on $\mathfrak{M}$ and $\lim_{n\to\infty} G_n = G$ uniformly on each compact subset of $\mathfrak{M}$. Now

$$\int_\gamma | G_n^*(\lambda) | d\omega(\lambda) = \frac{1}{2\pi} \int_0^{2\pi} | G_n^*(e^{i\theta}) | W(e^{i\theta}) e^{i\theta} d\theta.$$

But $W(e^{i\theta}) e^{i\theta}$ has a positive lower bound on $(0, 2\pi)$. Hence $\int_0^{2\pi} | G_n^*(e^{i\theta}) | d\theta < K'$, all n, by (1). Hence $\int_0^{2\pi} | G_n(re^{i\theta}) | d\theta < K''$, all n, by (d) of Lemma 4, $r' < r_1 < r < 1$. It follows that $\int_0^{2\pi} | G(re^{i\theta}) | d\theta < K''$, whence $G \varepsilon \mathfrak{H}'$. We claim that $G^*(\lambda) = G_0(\lambda)$ a. e.—$d\omega$.

Let now $r' < r < 1$. Set

$$U_n^1(z) = \frac{1}{2\pi i}\int_{|\tau|=r'} (\tau - z)^{-1} G_n(\tau)\, d\tau, \quad |z| > r'$$

$$U_n^2(z) = \frac{1}{2\pi i}\int_{|\tau|=r} (\tau - z)^{-1} G_n(\tau)\, d\tau, \quad |z| < r.$$

Then U_n^1 is analytic for $|z| > r'$, U_n^2 for $|z| < 1$, and

$$G_n(z) = U_n^2(z) - U_n^1(z), \quad r' < |z| < 1.$$

Clearly $U_n^{1*}(e^{i\theta}) = \lim_{z \to e^{i\theta}} U_n^1(z)$ exists for all θ. Also $G_n^*(e^{i\theta})$ exists a. e. Hence $U_n^{2*}(e^{i\theta})$ exists a. e. and $G_n^* = U_n^{2*} - U_n^{1*}$, a. e.

Now $G_n(\tau) \to G(\tau)$ uniformly on $|\tau| = r'$. Hence $U_n^1(z) \to U^1(z)$ uniformly for $r' < a \leqq |z| \leqq b < \infty$, and $U^{1*}(e^{i\theta})$ exists everywhere. Hence $U_n^2(z) \to U^2(z)$ uniformly in $r' < a \leqq |z| \leqq b < 1$, and $G(z) = U^2(z) - U^1(z)$. It follows that

(2) $$\int_0^{2\pi} | U^2(re^{i\theta}) |\, d\theta = O(1) \text{ as } r \to 1.$$

Also

(3) $$G^*(e^{i\theta}) = U^{2*}(e^{i\theta}) - U^{1*}(e^{i\theta}) \text{ a. e.}$$

Fix r, ϕ; $r < 1$, $0 \leqq \phi < 2\pi$. Set $g(\theta) = (1 - r^2)(1 + r^2 - 2r\cos(\theta - \phi))^{-1}$. Now

$$\frac{1}{2\pi}\int_0^{2\pi} U_n^{2*}(e^{i\theta}) g(\theta)\, d\theta = \frac{1}{2\pi}\int_0^{2\pi} (G_n^*(e^{i\theta}) + U_n^{1*}(e^{i\theta})) g(\theta)\, d\theta$$

and $U_n^{1*}(e^{i\theta}) \to U^{1*}(e^{i\theta})$ uniformly in $0 \leqq \theta < 2\pi$ and $G_n^*(\lambda)\, d\omega(\lambda)$ converges weakly to $G_0(\lambda)\, d\omega(\lambda)$. Hence

$$\lim_{n\to\infty} \frac{1}{2\pi}\int_0^{2\pi} U_n^{2*}(e^{i\theta}) g(\theta)\, d\theta = \frac{1}{2\pi}\int_0^{2\pi} (G_0(e^{i\theta}) + U^{1*}(e^{i\theta})) g(\theta)\, d\theta.$$

On the other hand,

$$U_n^2(re^{i\phi}) = \frac{1}{2\pi}\int_0^{2\pi} U_n^{2*}(e^{i\theta}) g(\theta)\, d\theta \to U^2(re^{i\phi}).$$

By (2) we get, since U^2 is analytic in $|z| < 1$,

$$U^2(re^{i\phi}) = \frac{1}{2\pi}\int_0^{2\pi} U^{2*}(e^{i\theta}) g(\theta)\, d\theta.$$

Hence

$$\int_0^{2\pi} U^{2*}(e^{i\theta}) g(\theta)\, d\theta = \int_0^{2\pi} (G_0(e^{i\theta}) + U^1(e^{i\theta})) g(\theta)\, d\theta,$$

or

$$0=\int_0^{2\pi}\{U^{2*}-G_0-U^{1*}\}(e^{i\theta})(1-r^2)(1+r^2-2r\cos(\theta-\phi))^{-1}d\theta.$$

This now holds for arbitrary r, ϕ. Since $U^{2*}-G_0-U^{1*}$ is summable on $(0,2\pi)$, we conclude

$$G_0(e^{i\theta})=U^{2*}(e^{i\theta})-U^{1*}(e^{i\theta})=G^*(e^{i\theta}),\text{ a. e.}$$

The conclusion now follows, as shown above.

THEOREM 1. *Let* $\mu_0\,\varepsilon\,\mathfrak{B}$ *and* $\int_\gamma f(\lambda)\,d\mu_0(\lambda)=0$ *for all* f *in* $\mathfrak{A}$. *Then there exists* $J\,\varepsilon\,\mathfrak{H}'$, $J(\zeta_0)=0$, *and constants* c_i *such that, setting*

$$L(\lambda)=J^*(\lambda)+\sum_{i=1}^{2p}c_iK_i(\lambda),$$

where K_i *are the functions of Lemma 2, we have* $d\mu_0(\lambda)=L(\lambda)\,d\omega(\lambda)$.

Proof. Let $\mathfrak{W}'$ be the vector-space obtained by adjoining to $\mathfrak{W}$ the measures $K_i(\lambda)\,d\omega(\lambda)$, $i=1,\cdots,2p$. Since $\mathfrak{W}$ is regularly closed, the same is true of $\mathfrak{W}'$. Our assertion amounts to the statement that $\mu_0\,\varepsilon\,\mathfrak{W}'$.

Suppose $\mu_0\not\varepsilon\,\mathfrak{W}'$. Since $\mathfrak{W}'$ is regularly closed, it follows by Banach's definition, that for some f_0 in $\boldsymbol{C}$

$$\int_\gamma f_0(\lambda)\,d\mu_0(\lambda)\neq 0 \tag{4}$$

$$\int_\gamma f_0(\lambda)\,d\mu(\lambda)=0\text{ if }\mu\,\varepsilon\,\mathfrak{W}'. \tag{5}$$

Let now $\bar{\mathfrak{A}}$ be the closure in $L^2(\gamma)$ of $\mathfrak{A}$. Then we can decompose f_0 as follows: $f_0=H+G$, $H\,\varepsilon\,\bar{\mathfrak{A}}$, G orthogonal to $\bar{\mathfrak{A}}$. Let $f\,\varepsilon\,\mathfrak{A}$. Then

$$\int_\gamma \bar{f}(\lambda)G(\lambda)\,d\omega(\lambda)=0. \tag{6}$$

Hence

$$\int_\gamma G(\lambda)\,d\omega(\lambda)=0. \tag{6'}$$

Let now $H_n\,\varepsilon\,\mathfrak{A}$, $H_n\to H$ in the norm of $L^2(\gamma)$. Then by the residue theorem

$$\frac{1}{2\pi i}\int_\gamma H_n(\lambda)(f(\lambda)-f(\zeta_0))\,d\omega(\lambda)=0,$$

whence

$$\int_\gamma H(\lambda)(f(\lambda)-f(\zeta_0))\,d\omega(\lambda)=0. \tag{7}$$

Also $(f(\lambda)-f(\zeta_0))d\omega(\lambda)\,\varepsilon\,\mathfrak{W}'$, whence by (5)

$$\int_\gamma f_0(\lambda)(f(\lambda)-f(\zeta_0))d\omega(\lambda)=0.$$

Hence from $f_0=H+G$ and (7),

$$\int_\gamma G(\lambda)(f(\lambda)-f(\zeta_0))d\omega(\lambda)=0.$$

By (6′), then,

$$\int_\gamma G(\lambda)f(\lambda)d\omega(\lambda)=0. \tag{8}$$

It follows from (6) and (8) that $\int_\gamma u(\lambda)G(\lambda)d\omega(\lambda)=0$ for all real continuous functions u on γ with $u=\operatorname{Re}f$, for some $f\,\varepsilon\,\mathfrak{A}$. As in the proof of Lemma 2, we get from this that for some b_ν, $G(\lambda)d\omega(\lambda)=\sum_{\nu=1}^{2p}b_\nu\Phi_\nu$ as functionals, and hence that

$$G(\lambda)d\omega(\lambda)=\sum_{i=1}^{2p}c_iK_i(\lambda)d\omega(\lambda) \tag{9}$$

where K_i are the functions constructed in Lemma 2. Now

$$\int_\gamma H(\lambda)K_\nu(\lambda)d\omega(\lambda)=0, \qquad \nu=1,\cdots,2p;$$

also $K_\nu(\lambda)d\omega(\lambda)\,\varepsilon\,\mathfrak{W}'$. Hence by (5),

$$\int_\gamma f_0(\lambda)K_\nu(\lambda)d\omega(\lambda)=0, \qquad \nu=1,\cdots,2p.$$

Hence

$$\int_\gamma G(\lambda)K_\nu(\lambda)d\omega(\lambda)=0, \qquad \nu=1,\cdots,2p. \tag{10}$$

By (9) and (10),

$$\int_\gamma |G(\lambda)|^2d\omega(\lambda)=\int_\gamma G(\lambda)\sum_{i=1}^{2p}\bar{c}_iK_i(\lambda)d\omega(\lambda)$$
$$=\sum_{i=1}^{2p}\bar{c}_i\int_\gamma G(\lambda)K_i(\lambda)d\omega(\lambda)=0.$$

Hence $G(\lambda)=0$ a. e. and so $f_0=H$ a. e.

Now consider $H_n\,\varepsilon\,\mathfrak{A}$, $H_n\to H$ in the norm of $L^2(\gamma)$. Then for $\zeta\,\varepsilon\,\mathfrak{M}$,

$$H_n(\zeta)=\frac{1}{2\pi i}\int_\gamma H_n(\lambda)W_\zeta(\lambda)d\lambda.$$

It then follows that

$$\frac{1}{2\pi i}\int_\gamma f_0(\lambda)W_\zeta(\lambda)\,d\lambda = \lim_{n\to\infty} H_n(\zeta)$$

is analytic in $\mathfrak{M}$. Also $\frac{1}{2\pi i}\int_\gamma f_0(\lambda)W_\zeta(\lambda)\,d\lambda$ has f_0 as continuous boundary value on γ. Hence $f_0 \in \mathfrak{A}$. Then $\int_\gamma f_0(\lambda)\,d\mu_0(\lambda) = 0$. This contradicts (4). Hence the assertion must be true.

3. Proof of Theorem 2. Let $\mathfrak{A}'$ be a closed subalgebra of $\boldsymbol{C}$ with $\mathfrak{A}' \neq \boldsymbol{C}$ and $\mathfrak{A} \subseteq \mathfrak{A}'$. Since $\mathfrak{A}'$ is a proper closed subspace of $\boldsymbol{C}$, a well-known theorem on Banach spaces guarantees the existence of a non-zero functional on $\boldsymbol{C}$ which annihilates $\mathfrak{A}'$. Thus there exists $\mu \in \mathfrak{B}$, $\mu \neq 0$ with $\int_\gamma g(\lambda)\,d\mu(\lambda) = 0$ if $g \in \mathfrak{A}'$. This holds in particular if $g \in \mathfrak{A}$. Hence by Theorem 1, $d\mu(\lambda) = L_0(\lambda)\,d\omega(\lambda)$ where L_0 is meromorphic on $\mathfrak{M}$ and analytic on $\mathfrak{M}$ except at the poles of the K_i. Hence L_0 is analytic except at the points $z_1,\cdots,z_k$ where W vanishes. Also, $\lim_{\zeta\to\lambda} L_0(\zeta)$ exists for a. a. λ on γ, if $\zeta \to \lambda$ within some sector, and this limit $\neq 0$ a. e. on γ.

Fix now $\phi \in \mathfrak{A}'$. We shall show $\phi \in \mathfrak{A}$. For if $f \in \mathfrak{A}$, $f(\lambda)\phi^m(\lambda) \in \mathfrak{A}'$ for $m = 1, 2, \cdots$ whence $\int_\gamma f(\lambda)\phi^m(\lambda)\,d\mu(\lambda) = 0$. Applying Theorem 1 to the measures $\phi^m(\lambda)\,d\mu(\lambda)$, we get $\phi^m(\lambda)\,d\mu(\lambda) = L_m(\lambda)\,d\omega(\lambda)$ where L_m has the same analyticity and boundary behavior as L_0. Hence $\phi^m(\lambda)L_0(\lambda) = L_m(\lambda)$ a. e. on γ. It follows that $(L_1(\lambda))^m = L_m(\lambda)\,(L_0(\lambda))^{m-1}$ a. e. on γ. On both sides we have non-tangential boundary values of functions analytic in the region $\mathfrak{M}_0$ obtained by deleting from $\mathfrak{M}$ the points $z_1,\cdots,z_k$. By a result of Lusin and Privaloff, [6], an analytic function possessing non-tangential boundary values on a set of positive measure is determined by these values. Hence $(L_1(\zeta))^m = L_m(\zeta)\,(L_0(\zeta))^{m-1}$ for ζ in $\mathfrak{M}_0$. Since this is true for all $m \geqq 1$, L_0 cannot have a zero at any point ζ' in $\mathfrak{M}_0$ of order α unless L_1 has at ζ' a zero of order $\geqq \alpha$. Hence $L_0^{-1}L_1$ is analytic in $\mathfrak{M}_0$. Also, since $\phi(\lambda)L_0(\lambda) = L_1(\lambda)$ a. e. on γ, ϕ is the non-tangential limit of $L_0^{-1}L_1$ a. e. on γ.

Set $T\phi(\zeta) = L_0^{-1}(\zeta)L_1(\zeta)$. The map $\phi \to T\phi$ then assigns to each ϕ in $\mathfrak{A}'$ an analytic function $T\phi$ on $\mathfrak{M}_0$ having boundary values $\phi(\lambda)$. By the theorem in [6] mentioned above, ϕ determines $T\phi$. Let now ϕ_1, ϕ_2 belong to $\mathfrak{A}'$. Then

$$\operatorname*{Lim}_{\zeta\to\lambda} T\phi_1(\zeta)\cdot T\phi_2(\zeta) = \phi_1(\lambda)\phi_2(\lambda)$$

and so $T(\phi_1 \cdot \phi_2) = T\phi_1 \cdot T\phi_2$. Similarly $T(\phi_1 + \phi_2) = T\phi_1 + T\phi_2$. Fix now z_0 in $\mathfrak{M}_0$. Then the map $\phi \to T\phi(z_0)$ is a multiplicative functional defined on $\mathfrak{A}'$. But a multiplicative functional on a Banach algebra is always bounded and has bound 1. Hence $| T\phi(z_0) | \leqq \| \phi \|$. Since z_0 is an arbitrary point in $\mathfrak{M}_0$, $T\phi$ is then bounded on $\mathfrak{M}_0$; hence $T\phi$ is analytic and bounded on $\mathfrak{M}$. Lemma 4 gives now that for ζ in $\mathfrak{M}$,

$$T\phi(\zeta) = \int_\gamma (T\phi)^*(\lambda)\, d\omega_\zeta(\lambda) = \int_\gamma \phi(\lambda)\, d\omega_\zeta(\lambda).$$

On the other hand the last integral represents a continuous function on $\mathfrak{M} + \gamma$ agreeing with $\phi(\lambda)$ on γ. Hence ϕ is in $\mathfrak{A}$, as asserted.

Hence $\mathfrak{A}' = \mathfrak{A}$, and so Theorem 2 is established.

4. (Added November 27, 1954.) Let now $\mathfrak{F}_1$, $\mathfrak{F}_2$ be Riemann surfaces, $\mathfrak{M}_1$, $\mathfrak{M}_2$ regions on them bounded by simple closed analytic curves γ_1, γ_2 with $\mathfrak{M}_i \cup \gamma_i$ compact, $i = 1, 2$. Let $\mathfrak{A}_i$ be the algebra of functions continuous on γ_i and extendable to be analytic on $\mathfrak{M}_i$, $i = 1, 2$. We assert:

THEOREM 3. *$\mathfrak{A}_1$ is isomorphic to $\mathfrak{A}_2$ as algebra if and only if $\mathfrak{M}_1$ is conformally equivalent to $\mathfrak{M}_2$.*

We need the following:

LEMMA. *If χ is a multiplicative functional on $\mathfrak{A}_i$, then there exists a point $p \,\varepsilon\, \mathfrak{M}_i \cup \gamma_i$ with $\chi(f) = f(p)$, all $f \,\varepsilon\, \mathfrak{A}_i$.*

Proof.[4] (We omit the subscript i from $\mathfrak{A}_i$, etc.) By the general representation theorem for bounded linear functions on spaces of continuous functions, there is a measure μ_0 on γ with

$$\chi(f) = \int_\gamma f(\lambda)\, d\mu_0(\lambda), \qquad f \,\varepsilon\, \mathfrak{A}.$$

Suppose now that the assertion of the Lemma is false. Then for each $p \,\varepsilon\, \mathfrak{M} \cup \gamma$ there exists $f_p \,\varepsilon\, \mathfrak{A}$ with $\chi(f_p) = 0$ and $f_p(p) \neq 0$.

Let $d\omega(\lambda)$ and $W(\zeta)$ have the same meaning as in the preceding sections. Let $\mathfrak{M}_0$ be the region obtained by deleting from $\mathfrak{M}$ the zeros of W.

Now for all $f \,\varepsilon\, \mathfrak{A}$ and p in $\mathfrak{M} \cup \gamma$

$$0 = \chi(f \cdot f_p) = \int_\gamma f(\lambda) f_p(\lambda)\, d\mu_0(\lambda).$$

Hence the measure $f_p(\lambda)\, d\mu_0(\lambda)$ annihilates $\mathfrak{A}$. By Theorem 1, then we can

[4] Cf. L. Carleson [7], Theorem 4, for a similar method of proof.

find a function L_p analytic on $\mathfrak{M}_0$ and with $L_p(\zeta)W(\zeta)$ regular on $\mathfrak{M}$, such that L_p has nontangential boundary-values $L_p(\lambda)$ for all λ in γ except for a set of ω-measure 0, and with $f_p(\lambda)\,d\mu_0(\lambda)=L_p(\lambda)\,d\omega(\lambda)$ as measures. Choose now p_1, p_2 distinct in $\mathfrak{M}\cup\gamma$. Then

$$f_{p_2}(\lambda)f_{p_1}(\lambda)\,d\mu_0(\lambda)=f_{p_2}(\lambda)L_{p_1}(\lambda)\,d\omega(\lambda)$$

and

$$f_{p_1}(\lambda)f_{p_2}(\lambda)\,d\mu_0(\lambda)=f_{p_1}(\lambda)L_{p_2}(\lambda)\,d\omega(\lambda)$$

whence $f_{p_2}\cdot L_{p_1}=f_{p_1}\cdot L_{p_2}$ a.e.—$d\omega$ on γ, whence by the result in [6] which we have quoted earlier, $f_{p_2}(\zeta)L_{p_1}(\zeta)=f_{p_1}(\zeta)L_{p_2}(\zeta)$ for all ζ in $\mathfrak{M}$.

Fix now p_0 in $\mathfrak{M}$ and set $F(\zeta)=f_{p_0}^{-1}(\zeta)L_{p_0}(\zeta)$. Since, for $q\in\mathfrak{M}_0$, L_q and f_q^{-1} are regular at q, we obtain that F is regular at q. Thus F is analytic on all $\mathfrak{M}_0$ and similarly we see that the covariant $F(\zeta)W(\zeta)$ is analytic on all of $\mathfrak{M}$.

Next, for each $q\in\gamma$, we choose an arc γ_q on γ with $|f_q(\lambda)|\geqq\delta_q$ for λ in γ_q, δ_q being a positive number. By the Heine-Borel theorem, some finite set of these arcs covers γ. We can hence get $\delta>0$ and a decomposition $\gamma=\bigcup_{i=1}^{n}\gamma_i$ where the γ_i are disjoint half-open arcs and for each i there is some q_i with $|f_{q_i}(\lambda)|\geqq\delta$ on γ_i.

Now, $F(\zeta)=f_{q_i}^{-1}(\zeta)L_{q_i}(\zeta)$ for $\zeta\in\mathfrak{M}$, whence

$$F(\lambda)=\lim_{\zeta\to\lambda}F(\zeta)=f_{q_i}^{-1}(\lambda)L_{q_i}(\lambda)\ \text{a.e. on}\ \gamma_i.$$

We now use annular coordinates r,θ: $r_0\leqq r\leqq 1$, $0\leqq\theta\leqq 2\pi$ in an annular subregion of $\mathfrak{M}$ bounded on one side by γ, with $r=1$ being the equation of γ.

Let $g\in\mathfrak{A}$. Then for each i

$$\int_{\gamma_i}g(\lambda)\,d\mu_0(\lambda)=\int_{\gamma_i}g(\lambda)f_{q_i}^{-1}(\lambda)L_{q_i}(\lambda)\,d\omega(\lambda)=\int_{\gamma_i}F(\lambda)g(\lambda)\,d\omega(\lambda).$$

Hence

$$\int_{\gamma}g(\lambda)\,d\mu_0(\lambda)=\int_{\gamma}g(\lambda)F(\lambda)\,(2\pi i)^{-1}W(\lambda)\,d\lambda.$$

Now if γ_ρ is the curve with equation: $r=\rho$, $\rho<1$,

$$\int_{\gamma_\rho}F(\zeta)g(\zeta)W(\zeta)\,d\zeta=0$$

by the residue theorem. Also

$$\lim_{\rho\to 1}\int_{\gamma_\rho} F(\zeta)g(\zeta)W(\zeta)d\zeta = \int_\gamma g(\lambda)F(\lambda)W(\lambda)d\lambda$$

due to the boundary behavior of the functions L_p and f_p. Hence

$$\chi(g) = \int_\gamma g(\lambda)\,d\mu_0(\lambda) = \tfrac{1}{2\pi i}\int_\gamma g(\lambda)F(\lambda)W(\lambda)d\lambda = 0.$$

This must hold for all $g \in \mathfrak{A}$, which is impossible. Hence the assertion of the Lemma must be true.

COROLLARY. *The space $\mathfrak{S}$ of multiplicative functionals on $\mathfrak{A}$ is homeomorphic to the set $\mathfrak{M} \cup \gamma$.*

Proof. By the Lemma, if $\chi \in \mathfrak{S}$, then there exists $p \in \mathfrak{M} \cup \gamma$ with $\chi(f) = f(p)$ for all $f \in \mathfrak{A}$. There cannot exist two distinct points p_1, p_2 with this property, for if $p_1 \neq p_2$ then for some f in $\mathfrak{A}$, $f(p_1) \neq f(p_2)$. Hence the map $\chi \to p$ takes $\mathfrak{S}$ into $\mathfrak{M} \cup \gamma$. It is obviously one-one and it is onto $\mathfrak{M} \cup \gamma$ since each p in $\mathfrak{M} \cup \gamma$ defines some multiplicative functional on $\mathfrak{A}$. Finally, the map is easily seen to be bicontinuous.

Proof of Theorem 3. Let τ be an algebraic isomorphism of $\mathfrak{A}_1$ onto $\mathfrak{A}_2$. Fix p in $\mathfrak{M}_2 \cup \gamma_2$. Map each f in $\mathfrak{A}_1$ into $\tau(f)(p)$. This map is a multiplicative functional on $\mathfrak{A}_1$, whence by the lemma there exists $\phi(p)$ in $\mathfrak{M}_1 \cup \gamma_1$ with $\tau(f)(p) = f(\phi(p))$ if $f \in \mathfrak{A}_1$. The function ϕ then maps $\mathfrak{M}_2 \cup \gamma_2$ onto $\mathfrak{M}_1 \cup \gamma_1$ in a one-one and bicontinuous fashion. It follows that ϕ maps $\mathfrak{M}_2$ homeomorphically onto $\mathfrak{M}_1$.

Fix p_0 in $\mathfrak{M}_2$ and f_0 in $\mathfrak{A}_1$ with f_0 locally simple at $\phi(p_0)$. Then for p in some neighborhood of p_0, $f_0(\phi(p)) = \tau(f_0)(p)$. Since f_0 and $\tau(f_0)$ are analytic functions and moreover f_0 is one-one in a neighborhood of $\phi(p_0)$, ϕ is analytic at p_0 as mapping from $\mathfrak{M}_2$ to $\mathfrak{M}_1$. This holds for each p_0 in $\mathfrak{M}_2$ and further ϕ is globally one-one. Hence ϕ provides a conformal map of $\mathfrak{M}_2$ onto $\mathfrak{M}_1$.

Conversely, suppose we are given a conformal map ϕ of $\mathfrak{M}_2$ on $\mathfrak{M}_1$. Classical results then give that ϕ is extendable to a homeomorphism of $\mathfrak{M}_2 \cup \gamma_2$ onto $\mathfrak{M}_1 \cup \gamma_1$. For each f in $\mathfrak{A}_1$ we can then define τf on $\mathfrak{M}_2 \cup \gamma_2$ as follows: $\tau f(p) = f(\phi(p))$, $p \in \mathfrak{M}_2 \cup \gamma_2$. Then $\tau f \in \mathfrak{A}_2$ and τ is an isomorphism from $\mathfrak{A}_1$ to $\mathfrak{A}_2$. This proves Theorem 3.

BROWN UNIVERSITY.

REFERENCES.

[1] J. Wermer, "On algebras of continuous functions," *Proceedings of the American Mathematical Society*, vol. 4 (1953), pp. 866-869.

[2] ———, "Algebras with two generators," *American Journal of Mathematics*, vol. 76 (1954), pp. 853-859.

[3] R. Nevanlinna, *Uniformisierung*, Springer Verlag, 1953.

[4] F. and M. Riesz, "Über Randwerte einer analytischen Funktion," *Quatrième Congrès des math. scandinaves*, 1916, pp. 27-44.

[5] S. Banach, *Opérations Linéaires*, Warszawa, 1932.

[6] N. N. Lusin and I. I. Privaloff, "Sur l'unicité et multiplicité des fonctions analytiques," *Annales de l'École Normale Supérieure*, ser. 3, vol. 42 (1925), pp. 143-191.

[7] L. Carleson, "On bounded analytic functions and closure problems," *Arkiv för Matematik*, 2, 12 (1952), pp. 283-291.

SOME RESULTS ON COHOMOTOPY GROUPS.*

By Franklin P. Peterson.[1]

1. Introduction. One of the central problems of topology is the computation of the set of homotopy classes of maps of a complex K into the n-sphere S^n. In 1936, Borsuk [2] showed that if the dimension of $K = N \leqq 2n - 2$, then this set admits a natural abelian group structure. In this case, this set is called the n-th cohomotopy group of K and denoted by $\pi^n(K)$. In 1949, Spanier [13] derived the basic properties of these groups and expressed the existing theorems on the structure of $\pi^n(K)$ by means of an exact sequence [13; p. 240]. These theorems are the Hopf theorem [7], which states that the natural homomorphism $\eta^n : \pi^n(K) \to H^n(K)$ ($=$ the n-th cohomology group of K) is an isomorphism for $n = N$ and is onto in case $n = N - 1$, and the Steenrod theorem [14], which computes the kernel of η^{N-1} and the image of η^{N-2}. Little more is known about the structure of $\pi^n(K)$.

In this paper, we shall derive further results concerning the structure of $\pi^n(K)$. First, $\pi^n(K)$ is finitely generated when K is finite. Second, $\pi^n(K)$ and $H^n(K)$ have the same rank. Third, the Hopf result is generalized by determining, for each prime p, a range of values of n for which η^n gives an isomorphism on the p-primary components of $\pi^n(K)$ and $H^n(K)$. Finally, the Steenrod result is generalized by giving a computation of the kernel and cokernel of η^n restricted to the p-primary components for a range of values of n where η^n is not an isomorphism. This computation is given in terms of the reduced p-th power operations of Steenrod [15]. In proving these results, we make use of a cohomotopy exact couple similar to that of Massey [8; part III] and of Serre's technique of "isomorphisms modulo a class of groups" [11].

In conclusion, I wish to express my warm appreciation to Professor N. E. Steenrod for his kind advice and encouragement. This paper is essentially Part I of a paper written under his direction and submitted as a dissertation to Princeton University.

* Received September 16, 1955.

[1] The author was a predoctoral National Science Foundation Fellow during the preparation of this paper.

2. Preliminaries. In this section, we recall the notions and notations which we need in order to state our main results.

We first recall the definition and elementary properties of cohomotopy groups [13]. Let K be a finite dimensional CW-complex [6], and let L be a subcomplex. Let $a:(K,L)\to(S^n,\text{pt.})$ be a continuous map, where S^n denotes the n-dimensional sphere and "pt." denotes any fixed point of S^n. Let $[a]$ denote the homotopy class of a. The set of all such homotopy classes has a natural abelian group structure defined on it if dimension $K=N\leq 2n-2$. We call this the n-th cohomotopy group of the CW-pair (K,L) and denote it by $\pi^n(K,L)$. A map $f:(K,L)\to(K',L')$ induces a homomorphism $f^{\#}:\pi^n(K',L')\to\pi^n(K,L)$ defined by $f^{\#}([a])=[af]$.

Let $\pi_r(X)$ denote the r-th homotopy group of the space X. The process of suspension induces a homomorphism $S_{\#}:\pi_r(S^n)\to\pi_{r+1}(S^{n+1})$ which is an isomorphism for $r<2n-1$ by the Freudenthan theorem [16]. We identify these groups under this isomorphism and denote the result by $Z_{(r-n)}$.

The homology theory best suited for our investigations is the cellular homology theory as described in [6]. We denote the n-th homology group of (K,L) with coefficients in G by $H_n(K,L;G)$ and the n-th cohomology group of (K,L) with coefficients in G by $H^n(K,L;G)$.

We denote the additive group of integers by Z, the group of integers mod n by Z_n, and the p-primary component of a group A by A_p, where the p-primary component of A is the subgroup of all elements of A whose orders are a power of the prime p. (The only exception to this notation is Z_p which denotes the integers mod p.) Let $\phi:A\to B$ be a homomorphism. We denote $\phi|A_p:A_p\to B_p$ by $\phi_{(p)}$. Furthermore, we denote the kernel of ϕ by $\operatorname{Ker}\phi$, the image of ϕ by $\operatorname{Im}\phi$, and the cokernel of ϕ $(=B/\operatorname{Im}\phi)$ by $\operatorname{Coker}\phi$. $A\otimes B$ and $\operatorname{Tor}(A,B)$ denote the tensor product and the torsion product respectively [4].

We now recall the notion of a class which was introduced by Serre [11]. A *class* $\mathscr{C}$ is a non-empty family of abelian groups such that

(I) if $0\to A\to B\to C\to 0$ is an exact sequence [5],
then $A\in\mathscr{C}$ and $C\in\mathscr{C}$ if and only if $B\in\mathscr{C}$.

In the applications, one of the following axioms is also assumed:

($\mathrm{II_A}$) if $A\in\mathscr{C}$ and $B\in\mathscr{C}$, then $A\otimes B\in\mathscr{C}$ and $\operatorname{Tor}(A,B)\in\mathscr{C}$, or

($\mathrm{II_B}$) if $A\in\mathscr{C}$ and B is arbitrary, then $A\otimes B\in\mathscr{C}$ and $\operatorname{Tor}(A,B)\in\mathscr{C}$.

The important examples of classes are $\mathscr{C}_0=$ the class consisting of the

0 group alone, $\mathscr{C}_T =$ the family of torsion groups, $\mathscr{C}_p =$ the family of torsion groups whose p-primary components are 0; $\mathscr{F} =$ the family of finitely generated groups, $\mathscr{C}_f =$ the family of finite groups, and $\mathscr{C}_{pf} =$ the family of finite groups whose p-primary components are 0. It is easily checked that $\mathscr{C}_0$, $\mathscr{C}_T$, and $\mathscr{C}_p$ satisfy axiom ($\mathrm{II_B}$), while $\mathscr{F}$, $\mathscr{C}_f$, and $\mathscr{C}_{pf}$ satisfy axiom ($\mathrm{II_A}$) but not ($\mathrm{II_B}$).

The notion of class was introduced by Serre to allow us to ignore systematically certain groups. With this in mind, we make the following definitions: a homomorphism $\phi: A \to B$ is a *$\mathscr{C}$-monomorphism* if $\operatorname{Ker} \phi \in \mathscr{C}$; it is a *$\mathscr{C}$-epimorphism* if $\operatorname{Coker} \phi \in \mathscr{C}$; it is a *$\mathscr{C}$-isomorphism* if both $\operatorname{Ker} \phi$ and $\operatorname{Coker} \phi \in \mathscr{C}$.

For any class $\mathscr{C}$, let $\alpha(\mathscr{C})$ denote the largest integer such that $Z_{(s)} \in \mathscr{C}$ for $0 < s < \alpha(\mathscr{C})$.

THEOREM 2.1. (a) *$Z_{(r)}$ is finite if $r > 0$,*

(b) $(Z_{(r)})_p = Z_p$ *if* $r = 2p - 3$

$\qquad = 0$ *otherwise for* $r < 4p - 5$,

(c) $\alpha(\mathscr{C}_T) = \infty$,

(d) $\alpha(\mathscr{C}_p) = 2p - 3$, *and*

(e) $\alpha(\mathscr{C}_0) = 1$.

Proof. (a) is a result of Serre [10]. (b) is a result of Serre [11]. (c) follows from (a). (d) follows from (a) and (b). It is well-known that $Z_{(1)} = \pi_{n+1}(S^n) = Z_2$, hence $\alpha(\mathscr{C}_0) = 1$.

3. The main results. The purpose of this section is to state our main results on the structure of cohomotopy groups. The remainder of this paper is devoted to the proofs of these results.

There is a natural homomorphism $\eta^r: \pi^r(K, L) \to H^r(K, L)$. η^r is defined as follows: let $a \in [a] \in \pi^r(K, L)$, and let u be a chosen generator of $H^r(S^r, \text{pt.})$. Then $\eta^r([a]) = a^*(u) \in H^r(K, L)$ (see Section 4 and [13; p. 234]). We study the relations between the cohomotopy groups and the cohomology groups using this homomorphism. The classical Hopf theorem [7] states that if K is an N-dimensional complex, then η^N is an isomorphism. Our first result extends this theorem modulo classes.

Let (K, L) be a CW-pair with dimension $K = N$ throughout the rest of this paper.

THEOREM 3.1. *Let $\mathscr{C}$ be a class satisfying condition* ($\mathrm{II_B}$) *of Section 2. Let $n > (N+1)/2$ be such that $H^r(K,L) \varepsilon \mathscr{C}$ for every $r > n$. Then η^r is a $\mathscr{C}$-isomorphism if $r > \mathrm{Max}((N+1)/2, n-\alpha(\mathscr{C}))$, and is a $\mathscr{C}$-epimorphism for $r = n-\alpha(\mathscr{C})$ in case $n-\alpha(\mathscr{C}) > (N+1)/2$.*

THEOREM 3.2. *Let $\mathscr{C}$ be a class satisfying condition* ($\mathrm{II_A}$) *of Section 2. Let $n > \mathrm{Max}((N+1)/2, N-\alpha(\mathscr{C}))$ be such that $H^r(K,L) \varepsilon \mathscr{C}$ for every $r > n$. Then η^r is a $\mathscr{C}$-isomorphism if $r \geqq n$, and is a $\mathscr{C}$-epimorphism for $r = n-1$.*

Theorems 3.1 and 3.2 solve a problem proposed by Steenrod in [9].

We have as an immediate corollary of Theorem 3.1:

COROLLARY 3.3. *Let K and L be two CW-complexes of dimensions M and N respectively, and let $f: L \to K$. Let $\mathscr{C}$ be a class satisfying condition* ($\mathrm{II_B}$) *of Section 2, and let $n > \mathrm{Max}((M+1)/2, (N+2)/2)$. Then the following two statements are equivalent:* (a) *$f^*: H^r(K) \to H^r(L)$ is a $\mathscr{C}$-isomorphism for $r > n$ and a $\mathscr{C}$-epimorphism for $r = n$, and* (b) *$f^\#: \pi^r(K) \to \pi^r(L)$ is a $\mathscr{C}$-isomorphism for $r > n$ and a $\mathscr{C}$-epimorphism for $r = n$.*

Proof. Replace f by a cellular approximation f' [6; p. 98]. By the mapping cylinder construction [6; p. 108], we may assume f' is an inclusion. Then (a) is true if and only if $H^r(K,L) \varepsilon \mathscr{C}$ for $r > n$ by the exact cohomology sequence of a pair. This is true if and only if $\pi^r(K,L) \varepsilon \mathscr{C}$ for $r > n$ by Theorem 3.1. However, this is true if and only if (b) is true by the exact cohomotopy sequence of a pair.

By specializing $\mathscr{C}$ to particular classes, we have the following four corollaries of Theorems 3.1 and 3.2.

COROLLARY 3.4. *Let $n > (N+1)/2$ be such that $H^r(K,L)$ is finitely generated for every $r > n$. Then $\pi^r(K,L)$ is finitely generated for $r > n$.*

Proof. This follows immediately from 3.2 by setting $\mathscr{C} = \mathscr{F}$ and noting 2.1 (a).

COROLLARY 3.5. *Let $n > (N+1)/2$ be such that $H^r(K,L)$ is finitely generated for every $r > n$. Then $\pi^r(K,L)$ and $H^r(K,L)$ have the same rank for $r > n$. Furthermore, if $r > (N+1)/2$ and $u \varepsilon H^r(K,L)$, then there is an integer $M \neq 0$ such that $Mu \varepsilon \mathrm{Im}\,\eta^r$.*

Proof. By 3.4, $\pi^r(K,L)$ and $H^r(K,L)$ are finitely generated for every $r > n$. Now apply 3.1 with $n = N$, $\mathscr{C} = \mathscr{C}_T$, and note that $\alpha(\mathscr{C}_T) = \infty$ by 2.1 (c). The conclusion that η^r is a $\mathscr{C}_T$-isomorphism for $r > n$ means that

$\pi^r(K,L)$ and $H^r(K,L)$ have the same rank. Furthermore, if for some $u \varepsilon H^r(K,L)$ there did not exist a non-zero integer M such that $Mu \varepsilon \operatorname{Im} \eta^r$, then $\operatorname{Coker} \eta^r \notin \mathscr{C}_T$.

The above is a result of Serre [11; p. 288].

COROLLARY 3.6. *Let $n > (N+1)/2 + 1$ be such that $H^r(K,L) = 0$ for every $r > n$. Then $\pi^r(K,L) = 0$ for $r > n$, η^n is an isomorphism, and η^{n-1} is an epimorphism.*

Proof. Set $\mathscr{C} = \mathscr{C}_0$, and use Theorem 3.1. Note that $\mathscr{C}_0$-isomorphism means regular isomorphism.

COROLLARY 3.7. *Let $n > (N+1)/2$ be such that $H^r(K,L) \varepsilon \mathscr{C}_p$ for every $r > n$. Then $\eta^r{}_{(p)}$ is an isomorphism for $r > \operatorname{Max}((N+1)/2, n-2p+3)$ and $\eta^{n-2p+3}{}_{(p)}$ is an epimorphism if $n-2p+3 > (N+1)/2$.*

Proof. Set $\mathscr{C} = \mathscr{C}_p$, and use Theorem 3.1. Note that ϕ being a $\mathscr{C}_p$-isomorphism implies that $\phi_{(p)}$ is an isomorphism.

For completeness, we state without proof a slight extension of a result of Steenrod [14] and Spanier [13; p. 240]. The theorem follows from an unpublished result of Adem (see [8; p. 263]) in a manner analogous to the way Theorem 3.9 follows from Theorem 6.2. Let $Sq^2 : H^{n-2}(K,L) \to H^n(K,L;Z_2)$ denote the Steenrod square [14]. Let $\Lambda : H^n(K,L;Z_2) \to \eta^{n-1}(K,L)$ be the homomorphism defined by Spanier [13; p. 238].

THEOREM 3.8. *Let $n > (N+1)/2 + 2$ be such that $H^r(K,L) = 0$ for every $r > n$. Then $\pi^r(K,L) = 0$ for $r > n$, η^n is an isomorphism, and the following sequence is exact:*

$$\pi^{n-2}(K,L) \xrightarrow{\eta^{n-2}} H^{n-2}(K,L) \xrightarrow{Sq^2} H^n(K,L;Z_2) \xrightarrow{\Lambda} \pi^{n-1}(K,L)$$
$$\xrightarrow{\eta^{n-1}} H^{n-1}(K,L) \longrightarrow 0.$$

Our final main result extends Corollary 3.7 in a manner analogous to the way Theorem 3.8 extends Corollary 3.6. Let

$$\mathscr{P}^1 : H^s(K,L) \to H^{s+2p-2}(K,L;Z_p)$$

denote the first Steenrod reduced p-th power operation [15].

$$\mu : H^{r+2p-3}(K,L;Z_p) \to \pi^r(K,L)_p$$

is a homomorphism to be defined in Section 6. In Section 6 we prove the following:

THEOREM 3.9. *Let $n > (N+1)/2$ be such that $H^r(K,L) \varepsilon \mathcal{C}_p$ for every $r > n$. Then $\eta^r{}_{(p)}$ is an isomorphism for $r > \mathrm{Max}((N+1)/2, n-2p+3)$ and the following sequences are exact:*

$$H^{r-1}(K,L) \xrightarrow{\mathcal{P}^1} H^{r+2p-3}(K,L;Z_p) \xrightarrow{\mu} \pi^r(K,L)_p \xrightarrow{\eta^r{}_{(p)}} H^r(K,L)_p$$

$$\xrightarrow{\mathcal{P}^1{}_{(p)}} H^{r+2p-2}(K,L;Z_p) \quad \textit{for } r > \mathrm{Max}((N+1)/2, n-4p+5)$$

and

$$\eta^{n-4p+5}(K,L)_p \xrightarrow{\eta^{n-4p+5}{}_{(p)}} H^{n-4p+5}(K,L)_p \xrightarrow{\mathcal{P}^1{}_{(p)}} H^{n-2p+3}(K,L;Z_p)$$

if $n-4p+5 > (N+1)/2+1$.

For $r > n-4p+5$, note that Theorem 3.9 computes the kernel and the cokernel of $\eta^r{}_{(p)}$ in terms of the cohomology groups of (K,L) and the first reduced p-th powers in $H^*(K,L)$.

4. The cohomotopy exact couple. The proofs of our main theorems are based on a cohomotopy exact couple of the pair (K,L) similar to the one studied by Massey [8; part III]. Since it differs from Massey's cohomotopy exact couple, we describe it in detail.

Let (K,L) be a CW-pair with dimension $K = N$. Let z be the least integer $> (N+1)/2$; i.e. z is the least integer n for which $\pi^n(K,L)$ has a natural group structure. Let K^s denote the union of L with the s-dimensional skeleton of K. Our exact couple is based on the exact cohomotopy sequence of the triple (K,K^s,K^{s-1}):

$$\pi^z(K,K^s) \xrightarrow{j} \pi^z(K,K^{s-1}) \to \cdots \to \pi^r(K,K^s) \xrightarrow{j} \pi^r(K,K^{s-1})$$

$$\xrightarrow{i} \pi^r(K^s,K^{s-1}) \xrightarrow{\Delta} \pi^{r+1}(K,K^s) \to \cdots,$$

where i and j are the homomorphisms induced by the inclusions

$$(K^s,K^{s-1}) \to (K,K^{s-1}) \text{ and } (K,K^{s-1}) \to (K,K^s)$$

respectively and Δ is the coboundary operator of the triple (K,K^s,K^{s-1}) (see [13; p. 229] for the definition of Δ and the proof of exactness).

For notational convenience, we set

$$A^{r,s} = \pi^r(K, K^s) \text{ for } r \geqq z,$$

$$C^{r,s} = \pi^r(K^s, K^{s-1}) \text{ for } r \geqq z.$$

Also let

$$j^{r,s}: A^{r,s} \to A^{r,s-1},$$

$$i^{r,s}: A^{r,s-1} \to C^{r,s}, \text{ and}$$

$$\Delta^{r,s}: C^{r,s} \to A^{r+1,s}$$

be the appropriate j, i, or Δ for $r \geqq z$. In order to extend the above sequence to an exact sequence extending indefinitely in both directions, we set

$$C^{z-1,s} = \operatorname{Ker} j^{z,s} \text{ for } r < z-1;$$

$$C^{r,s} = 0 \text{ for } r < z-1,$$

$$A^{r,s} = 0 \text{ for } r < z,$$

$\Delta^{z-1,s}: C^{z-1,s} \to A^{z,s}$ to be the inclusion, and the remaining homomorphisms i, j, and Δ to be zero.

The indices on the homomorphisms i, j, and Δ are determined by their domains, and thus we omit them whenever possible.

The groups and homomorphisms defined above fit together in a lattice as in Figure 1. Any path in Figure 1 which moves downward and to the right in a zig-zag pattern traces out an exact sequence. This follows immediately from the exact sequence of the triple (K, K^s, K^{s-1}) and our definitions.

$$\begin{tikzcd}
 & & \vdots \arrow[d] & & \vdots \arrow[d] & \\
 & \cdots \arrow[r] & A^{r,s} \arrow[r, "i"] \arrow[d, "j"] & C^{r,s+1} \arrow[r, "\Delta"] & A^{r+1,s+1} \arrow[r] \arrow[d, "j"] & \cdots \\
\cdots \arrow[r] & C^{r-1,s-1} \arrow[r, "\Delta"] & A^{r,s-1} \arrow[r, "i"] \arrow[d, "j"] & C^{r,s} \arrow[r, "\Delta"] & A^{r+1,s} \arrow[r] \arrow[d, "j"] & \cdots \\
 & \cdots \arrow[r] & A^{r,s-2} \arrow[r, "i"] \arrow[d] & C^{r,s-1} \arrow[r, "\Delta"] & A^{r+1,s-1} \arrow[r] \arrow[d] & \cdots \\
 & & \vdots & & \vdots &
\end{tikzcd}$$

Figure 1.

We now compute some of the groups of the cohomotopy exact couple. The following groups are obviously 0 for $m \geqq 1$: $C^{r,r-m}$, $C^{N+m,s}$, $A^{N+m,s}$, $A^{r,N+m-1}$, and $C^{r,N+m}$. Also $j: A^{r,r-m} \to A^{r,r-m-1}$ is an isomorphism for $m \geqq 2$ and is onto for $m = 1$. This follows because

$$C^{r-1,r-m} \to A^{r,r-m} \to A^{r,r-m-1} \to C^{r,r-m}$$

is exact and for $m \geqq 1$, $C^{r,r-m} = 0$. Hence $A^{r,r-2} \approx A^{r,-1} = \pi^r(K, L)$.

Define a homomorphism

$$\psi\colon \pi^r(K^s, K^{s-1}) \to C^s(K, L; \pi_s(S^r, \mathrm{pt.}))$$

as follows. The cellular homology and cohomology theory is based on $\pi_s(K^s, K^{s-1})$ as the group of chains in dimension s [7]. Hence

$$C^s(K, L; \pi_s(S^r, \mathrm{pt.})) = \mathrm{Hom}(\pi_s(K^s, K^{s-1}), \pi_s(S^r, \mathrm{pt.})),$$

where $\mathrm{Hom}(A, B)$ denotes the group of all homomorphisms from A to B. Let $[b] \in \pi_s(K^s, K^{s-1})$, $[a] \in \pi^r(K^s, K^{s-1})$, then define

$$\psi([a])([b]) = [ab] \in \pi_s(S^r, \mathrm{pt.}).$$

It is shown in [13; p. 222] that ψ has the following properties for $r \geqq z$ and $s \leqq N$:

1) ψ is an isomorphism,

2) ψ is natural with respect to cellular maps $f\colon (K, L) \to (K', L')$, and

3) the following diagram is commutative:

$$\begin{array}{ccc}
\pi^r(K^s, K^{s-1}) & \xrightarrow{i\Delta} & \pi^{r+1}(K^{s+1}, K^s) \\
\downarrow{\scriptstyle\psi} & & \downarrow{\scriptstyle\psi} \\
C^s(K, L; \pi_s(S^r, \mathrm{pt.})) & \xrightarrow{S_\#\delta} & C^{s+1}(K, L; \pi_{s+1}(S^{r+1}, \mathrm{pt.})),
\end{array}$$

where δ is the coboundary homomorphism

$$C^s(K, L; \pi_s(S^r, \mathrm{pt.})) \to C^{s+1}(K, L; \pi_s(S^r, \mathrm{pt.}))$$

and $S_\#$ is the homomorphism

$$C^{s+1}(K, L; \pi_s(S^r, \mathrm{pt.})) \to C^{s+1}(K, L; \pi_{s+1}(S^{r+1}, \mathrm{pt.}))$$

induced by suspending the coefficient group. Under our assumptions on r and s, $S_\#$ is an isomorphism on the coefficients, and we may write the commutative diagram as

$$\begin{array}{ccc}
\pi^r(K^s, K^{s-1}) & \xrightarrow{i\Delta} & \pi^{r+1}(K^{s+1}, K^s) \\
\downarrow{\scriptstyle\psi} & & \downarrow{\scriptstyle\psi} \\
C^s(K, L; Z_{(s-r)}) & \xrightarrow{\delta} & C^{s+1}(K, L; Z_{(s-r)})
\end{array}$$

(see Section 2 for the definition of $Z_{(s-r)}$).

We now describe the first derived cohomotopy exact couple (see [8; part I] for the prescise definitions and the proof that this is an exact couple). Define

$$\mathscr{H}^{r,s} = H(C^{r,s}) = \operatorname{Ker}(i^{r+1,s+1}\Delta^{r,s})/\operatorname{Im}(i^{r,s}\Delta^{r-1,s-1}),$$

$$\Gamma^{r,s} = \operatorname{Im} j^{r,s},$$

$$j'^{r,s}\colon \Gamma^{r,s} \to \Gamma^{r,s-1} \text{ by } j^{r,s-1},$$

$$i'^{r,s}\colon \Gamma^{r,s-1} \to \mathscr{H}^{r,s} \text{ by } i^{r,s}(j^{r,s-1})^{-1}, \text{ and}$$

$$\Delta'^{r,s}\colon \mathscr{H}^{r,s} \to \Gamma^{r+1,s+1} \text{ by } \Delta^{r,s}.$$

These groups and homomorphisms fit together in a lattice as in Figure 2.

From the remarks above, the following groups are obviously 0 for $m \geqq 1$: $\mathscr{H}^{r,r-m}$, $\mathscr{H}^{N+m,s}$, $\Gamma^{N+m,s}$, $\Gamma^{r,N+m-1}$, and $\mathscr{H}^{r,N+m}$. Also

$$\Gamma^{r,r-1} \approx \Gamma^{r,r-2} \approx \cdots \approx \pi^r(K,L) \quad \text{for } r \geqq z.$$

Furthermore, $\mathscr{H}^{r,s} \approx H^s(K,L;Z_{(s-r)})$ for $r \geqq z+1$ by the above identifications and definitions of $\mathscr{H}^{r,s}$.

$$\begin{tikzcd}
 & & \vdots \arrow[d] & & \vdots \arrow[d] & \\
 & \cdots \arrow[r] & \Gamma^{r,s} \arrow[r, "i'"] \arrow[d, "j'"] & \mathscr{H}^{r,s+1} \arrow[r, "\Delta'"] & \Gamma^{r+1,s+2} \arrow[r] \arrow[d, "j'"] & \cdots \\
\cdots \arrow[r] & \mathscr{H}^{r-1,s-2} \arrow[r, "\Delta'"] & \Gamma^{r,s-1} \arrow[r, "i'"] \arrow[d, "j'"] & \mathscr{H}^{r,s} \arrow[r, "\Delta'"] & \Gamma^{r+1,s+1} \arrow[r] \arrow[d, "j'"] & \cdots \\
 & \cdots \arrow[r] & \Gamma^{r,s-2} \arrow[r, "i'"] \arrow[d] & \mathscr{H}^{r,s-1} \arrow[r, "\Delta'"] & \Gamma^{r+1,s} \arrow[r] \arrow[d] & \cdots \\
 & & \vdots & & \vdots &
\end{tikzcd}$$

Figure 2.

Under these identifications, $i'^{r,r}\colon \pi^r(K,L) \to H^r(K,L)$ is defined geometrically as follows. Let $[a] \varepsilon \pi^r(K,L)$. We may assume that $a\colon (K,L) \to (S^r, \text{pt.})$ is such that $a(K^{r-1}) = \text{pt.}$ Restrict a to a map $a'\colon (K^r, K^{r-1}) \to (S^r, \text{pt.})$, then $[a']$ is an r-cochain of (K,L) which is a cocycle. Its cohomology class is $i'^{r,r}([a])$. We denote the homomorphism $i'^{r,r}$ by η^r. It is easily checked that this gives the same definition as given in Section 3.

We recall for reference later the definition of a cohomology operation θ of type $(n,q;A,B)$. θ is a function $\theta\colon H^n(K,L;A) \to H^q(K,L;B)$, defined for every CW-pair (K,L), such that if $f\colon (K,L) \to (K',L')$, then $f^*\theta = \theta f^*\colon H^n(K',L';A) \to H^q(K,L;B)$. A theorem of Serre [12; p. 220] states that the cohomology operations of a given type $(n,q;A,B)$ are in 1-1 correspondence with the elements of $H^q(A,n;B) = H^p(K(A,n);B)$, where $K(A,n)$ is an Eilenberg-MacLane complex (see [12] for details). This theorem holds if θ is only defined for every CW-pair (K,L) with dimension $K \leqq N$ and $N \geqq q+1$.

Let $f\colon (K,L)\to(K',L')$ be a cellular map; i.e. $f(K^n)\subset K'^n$. Then f induces homomorphisms

$$f^{\#}\colon \pi^r(K'^s,K'^{s-1})\to\pi^r(K^s,K^{s-1}) \text{ and } f^{\#}\colon \pi^r(K',K'^s)\to\pi^r(K,K^s)$$

which commute with i, j, and Δ when $r\geqq z, z'$, where $z'=$ the least integer $>$ (dimension $K'+1)/2$. Hence f induces a homomorphism of the cohomotopy exact couple of (K',L') into the cohomotopy exact couple of (K,L), and hence a homomorphism of the first derived cohomotopy exact couples. Thus this induced homomorphism f^* commutes with $i'\Delta'\colon \mathfrak{H}^{r,s}\to\mathfrak{H}^{r+1,s+2}$. Hence, $i'\Delta'$ is a cohomology operation because any two cellular approximations f_1 and f_2 to an arbitrary map $g\colon (K,L)\to(K',L')$ induce the same homomorphism g^* on $\mathfrak{H}^{r,s}$ and $\mathfrak{H}^{r+1,s+2}$ (see [6; p. 98] for the definition and properties of cellular approximations).

5. Proof of the Hopf theorem mod $\mathcal{C}$. In this section we give the proofs of Theorems 3.1 and 3.2.

Proof of 3.1. The proof is based on the first derived cohomotopy exact couple.[2] In order to prove that η^r is a $\mathcal{C}$-isomorphism, it suffices to show that $\Gamma^{r,r}\,\varepsilon\,\mathcal{C}$ for $r>\operatorname{Max}((N+1)/2, n-\alpha(\mathcal{C}))$ because

$$\Gamma^{r,r}\xrightarrow{j'}\pi^r(K,L)\xrightarrow{\eta^r}H^r(K,L)\xrightarrow{\Delta'}\Gamma^{r+1,r+1}$$

is an exact sequence. Again by exactness (and the fact that $\Gamma^{r,N}=0\,\varepsilon\,\mathcal{C}$), it suffices to prove that $H^{r+1}(K,L;Z_{(1)})\,\varepsilon\,\mathcal{C},\cdots,H^N(K,L;Z_{(N-r)})\,\varepsilon\,\mathcal{C}$ for $r>\operatorname{Max}((N+1)/2, n-\alpha(\mathcal{C}))$. Now $n-r<n-(n-\alpha(\mathcal{C}))=\alpha(\mathcal{C})$, hence $Z_{(1)}\,\varepsilon\,\mathcal{C},\cdots,Z_{(n-r)}\,\varepsilon\,\mathcal{C}$ by definition of $\alpha(\mathcal{C})$. Since $Z_{(s)}$ is finitely generated by 2.1 (a), we may use Theorem 2.3A:[3]

$$0\to H^{r+s}(K,L)\otimes Z_{(s)}\xrightarrow{\alpha}H^{r+s}(K,L;Z_{(s)})\xrightarrow{\beta}\operatorname{Tor}(H^{r+s+1}(K,L),Z_{(s)})\to 0$$

is an exact sequence. $Z_{(s)}\,\varepsilon\,\mathcal{C}$ for $s\leqq n-r$, hence $H^{r+s}(K,L;Z_{(s)})\,\varepsilon\,\mathcal{C}$ for $s\leqq n-r$ by properties (I) and (II$_B$) of classes. For $s>n-r$,

$$H^{r+s}(K,L)\,\varepsilon\,\mathcal{C} \text{ and } H^{r+s+1}(K,L)\,\varepsilon\,\mathcal{C}$$

by hypothesis, and again by the above exact sequence $H^{r+s}(K,L;Z_{(s)})\,\varepsilon\,\mathcal{C}$ for $s>n-r$. This completes the proof.

[2] It is suggested that the reader draw a diagram of the relevant portion of the first derived cohomotopy exact couple to facilitate following the proof.

[3] 2.3A stands for Theorem 2.3 of the appendix.

Proof of 3.2. This proof is very similar to that of 3.1. Again we need to show that $H^{r+1}(K,L;Z_{(1)}) \in \mathscr{C}, \cdots, H^N(K,L;Z_{(N-r)}) \in \mathscr{C}$ for $r \geqq n$. We use the above exact sequence. Since $H^{r+s}(K,L) \in \mathscr{C}$ for $s \geqq 1$ and $Z_{(s)} \in \mathscr{C}$ for $s \leqq N-n < N-(N-\alpha(\mathscr{C})) = \alpha(\mathscr{C})$, by conditions ($\mathrm{II}_A$) and (I) on classes the result is proven.

6. Proof of Theorem 3.9. Corollary 3.7 gives us the range of values of r for which $\eta^r{}_{(p)}$ is an isomorphism. The first open problems arising are determining the kernel of $\eta^{n-2p+3}{}_{(p)}$ and the image of $\eta^{n-2p+2}{}_{(p)}$ from the cohomology structure of the pair (K,L). This and more is achieved by the exact sequence of 3.9. In order to prove 3.9, we first prove a result which introduces the Steenrod reduced p-th power operations into the first derived cohomotopy exact couple.

These results lead us to the next open problem; namely, determining the kernel of $\eta^{n-4p+5}{}_{(p)}$ from the cohomology structure of the pair (K,L). This requires a further study of secondary (and higher order) cohomology operations. Some results on the nature of these operations, with applications to the above, have been obtained by the author and will appear at a later date.

LEMMA 6.1. *In the first derived cohomotopy exact couple,* $\Gamma^{r+1,r+s}$ *is a torsion group for* $s \geqq 1$. *Furthermore,* $j'_{(p)}: \Gamma^{r+1,r+s}{}_p \to \Gamma^{r+1,r+s-1}{}_p$ *is an isomorphism for* $2 \leqq s < 2p-2$.

Proof. The following exact sequence is part of the first derived cohomotopy exact couple:

$$H^{r+s-1}(K,L;Z_{(s-1)}) \xrightarrow{\Delta'} \Gamma^{r+1,r+s} \xrightarrow{j'} \Gamma^{r+1,r+s-1} \xrightarrow{i'} H^{r+s}(K,L;Z_{(s-1)}).$$

By 2.1 (a), $Z_{(s-1)} \in \mathscr{C}_T$ for $s \geqq 2$, hence $H^{r+s-1}(K,L;Z_{(s-1)}) \in \mathscr{C}_T$ for $s \geqq 2$. Also, $\Gamma^{r+1,N} = 0 \in \mathscr{C}_T$, therefore by induction, $\Gamma^{r+1,r+s-1} \in \mathscr{C}_T$ for $s \geqq 2$. Similarily, $Z_{(s-1)} \in \mathscr{C}_p$ for $2 \leqq s < 2p-2$ by 2.1 (b), hence $H^{r+s-1}(K,L;Z_{(s-1)}) \in \mathscr{C}_p$ and $H^{r+s}(K,L;Z_{(s-1)}) \in \mathscr{C}_p$ for $2 \leqq s < 2p-2$. Therefore, $j'_{(p)}: \Gamma^{r+1,r+s}{}_p \to \Gamma^{r+1,r+s-1}{}_p$ is an isomorphism for $2 \leqq s < 2p-2$. This completes the proof.

Let $r > (N+1)/2$. We define

$$d: H^r(K,L) \to H^{r+2p-2}(K,L;Z_{(2p-3)})_p = H^{r+2p-2}(K,L;Z_p)$$

by

$$d(u) = i'^{r+1,r+2p-2}{}_{(p)}(j'^{r+1,r+2p-3}{}_{(p)})^{-1} \cdots (j'^{r+1,r+2}{}_{(p)})^{-1} P\Delta'^{r,r}(u),$$

where $P: \Gamma^{r+1,r+1} \to \Gamma^{r+1,r+1}{}_p$ is the natural projection onto the p-primary component (P is naturally defined because $\Gamma^{r+1,r+1}$ is a torsion group).

THEOREM 6.2. *In the first derived cohomotopy exact couple, the homomorphism* $d\colon H^r(K,L) \to H^{r+2p-2}(K,L;Z_p)$ *is equal to* $\beta\mathcal{P}^1$, *where* $\beta \not\equiv 0$ $(\operatorname{mod} p)$.

Proof. Since all the homomorphisms in the definition of d are natural with respect to cellular maps $f\colon (K,L) \to (K',L')$, d is also natural with respect to such maps and hence with respect to all maps (as in Section 4). Thus d is a cohomology operation and d corresponds to an element of $H^{r+2p-2}(Z,r;Z_p) \approx Z_p$ (see the calculations of Cartan [3]). Therefore $d = \beta\mathcal{P}^1$, and it suffices to exhibit a complex K for which $d \neq 0$ for then $\beta \not\equiv 0$ $(\operatorname{mod} p)$.

Let $(K,L) = (M,x_0)$, where $M = S^r \cup e^{r+2p-2}$, the cell e^{r+2p-2} being attached to S^r by a non-zero element of $\pi_{r+2p-3}(S^r)_p = Z_p$. M has the property that $\mathcal{P}^1(u) \neq 0$, where u is a generator of $H^r(M,x_0)$ (this is a result of Borel and Serre [1; p. 425]). Assume $d = 0$ for this complex. Since $H^s(M,x_0;G) = 0$ unless $s = r$ or $s = r + 2p - 2$,

$$\Gamma^{r+1,r+1} \approx H^{r+2p-2}(M,x_0;Z_{(2p-3)}),$$

and $d(u) = P\Delta'^{r,r}(u)$ for $u \in H^r(M,x_0)$. Let u generate $H^r(M,x_0) \approx Z$. $\Delta'^{r,r}(u)$ is an element of finite order, hence there is an integer $D \equiv 1$ $(\operatorname{mod} p)$ such that $P\Delta'^{r,r}(u) = D\Delta'^{r,r}(u) = \Delta'^{r,r}(Du)$. Since $d = 0$, $\Delta'^{r,r}(Du) = 0$, and by exactness, $Du = \eta^r([a])$, where $a\colon (M,x_0) \to (S^r, \text{pt.})$. $a \mid S^r\colon (S^r,x_0) \to (S^r,\text{pt.})$ is a map of degree D, and hence $a^*(u') = Du$, where u' is a generator of $H^r(S^r,\text{pt.})$. Thus

$$0 \neq D\mathcal{P}^1(u) = \mathcal{P}^1(Du) = \mathcal{P}^1(a^*(u')) = a^*(\mathcal{P}^1(u')) = 0$$

because $D \equiv 1$ $(\operatorname{mod} p)$ and $\mathcal{P}^1(u') = 0$ in S^r. This is a contradiction, and hence $d \neq 0$. This completes the proof.

Using a more computational proof, it can be shown that $\beta \equiv 1$ $(\operatorname{mod} p)$. However, to prove Theorem 3.9, it is not necessary to know that $\beta \equiv 1$ $(\operatorname{mod} p)$ because $\operatorname{Im}(\beta\mathcal{P}^1) = \operatorname{Im}\mathcal{P}^1$ and $\operatorname{Ker}(\beta\mathcal{P}^1) = \operatorname{Ker}\mathcal{P}^1$ as long as $\beta \not\equiv 0$ $(\operatorname{mod} p)$. Using this remark, we now prove 3.9.

Proof of 3.9. The proof is based on the first derived cohomotopy exact couple.[2] By hypothesis, $r + 4p - 5 > n$, and hence $H^{r+4p-5}(K,L;Z_{(4p-5)}) \in \mathcal{C}_p$. It follows that $\Gamma^{r,r+2p-3} \in \mathcal{C}_p$ and $\Gamma^{r+1,r+2p-2} \in \mathcal{C}_p$. Hence

$$\Gamma^{r,r}{}_p \approx H^{r+2p-3}(K,L;Z_{(2p-3)})_p = H^{r+2p-3}(K,L;Z_p)$$

under the isomorphism $i'^{r,r+2p-3}{}_{(p)}(j'^{r,r+2p-4}{}_{(p)})^{-1} \cdots (j'^{r,r}{}_{(p)})^{-1}$. Furthermore,

$\operatorname{Im}(P\Delta'^{r-1,r-1}) = (\operatorname{Im}\Delta'^{r-1,r-1}) \cap \Gamma^{r,r}{}_p$. Similar remarks for $\Delta'^{r,r}$ and $\Gamma^{r+1,r+1}$ hold. From the exact couple, we obtain the exact sequence

$$H^{r-1}(K,L) \xrightarrow{P\Delta'^{r-1,r-1}} \Gamma^{r,r}{}_p \xrightarrow{j'^{r,r}{}_{(p)}} \pi^r(K,L)_p \xrightarrow{\eta^r{}_{(p)}} H^r(K,L)_p$$
$$\xrightarrow{\Delta'^{r,r}{}_{(p)}} \Gamma^{r+1,r+1}{}_p.$$

Using Theorem 6.2 and the above isomorphisms, we obtain the exact sequence of 3.9 with μ defined by

$$\mu = j'^{r,r}{}_{(p)} \cdots j'^{r,r+2p-4}{}_{(p)}\,(i'^{r,r+2p-3}{}_{(p)})^{-1}:$$
$$H^{r+2p-3}(K,L;Z_p) \to \pi^r(K,L)_p.$$

For $r = n - 4p + 5$, we have only the statement on the cokernel of $\eta^r{}_{(p)}$.

Appendix.

1. Results from Eilenberg and Steenrod [5]. The purpose of this appendix is to discuss the universal coefficient theorem for cohomology.

THEOREM 1.1. *Let K be a chain complex composed of free groups. For an arbitrary abelian group G, the following sequences are exact and split:*

$$0 \to H_r(K)\otimes G \xrightarrow{\alpha} H_r(K;G) \xrightarrow{\beta} \operatorname{Tor}(H_{r-1}(K),G) \to 0 \quad \textit{and}$$

$$0 \to \operatorname{Ext}(H_{r-1}(K),G) \xrightarrow{\beta} H^r(K;G) \xrightarrow{\alpha} \operatorname{Hom}(H_r(K),G) \to 0.$$

These exact sequences are natural with respect to chain maps $f\colon K \to K'$ and homomorphisms $\phi\colon G \to H$.

Proof. See exercise G-3 in [5; Chapt. V].

THEOREM 1.2. *Let K be a chain complex composed of finitely generated free groups. For an arbitrary abelian group G, the following sequences are exact and split:*

$$(*) \quad 0 \to H^r(K)\otimes G \xrightarrow{\alpha} H^r(K;G) \xrightarrow{\beta} \operatorname{Tor}(H^{r+1}(K),G) \to 0 \quad \textit{and}$$

$$0 \to \operatorname{Ext}(H^{r+1}(K),G) \xrightarrow{\beta} H_r(K;G) \xrightarrow{\alpha} \operatorname{Hom}(H^r(K),G) \to 0.$$

These exact sequences are natural with respect to chain maps $f\colon K \to K'$ and homomorphisms $\phi\colon G \to H$.

Proof. See exercises F-3, F-4, and G-3 in [5; Chapt. V].

Let $0 \to G' \xrightarrow{\phi} G \xrightarrow{\psi} G'' \to 0$ be an exact coefficient sequence.

$$\delta_*: H^r(K; G'') \to H^{r+1}(K; G')$$

is defined in [5; p. 158]. Define the sequence corresponding to this exact coefficient sequence to be the following sequence of groups and homomorphisms:

$$\cdots \to H^r(K; G') \xrightarrow{\phi_*} H^r(K; G) \xrightarrow{\psi_*} H^r(K; G'') \xrightarrow{\delta_*} H^{r+1}(K; G') \to \cdots.$$

THEOREM 1.3. *Let K be a chain complex composed of free groups. Then the sequence corresponding to the above exact coefficient sequence is exact. This exact sequence is natural with respect to chain maps $f: K \to K'$ and homomorphisms of one exact coefficient sequence into another.*

Proof. See exercise C-3 in [5; Chapt. V].

2. A new universal coefficient theorem. The sequence (*) of 1.2A does not necessarily hold if K is not finitely generated. We now prove a similar theorem by assuming G is finitely generated with K arbitrary.

LEMMA 2.1. *If G is finitely generated and free, then $\alpha: H^r(K) \otimes G \to H^r(K; G)$ is an isomorphism.*

Proof. α is obviously an isomorphism in case $G = Z$. Furthermore, the functors $H^r(K) \otimes G$ and $H^r(K; G)$ are additive with respect to G and thus commute with finite direct sums [4]. Hence α is an isomorphism if G is finitely generated and free.

THEOREM 2.2. *Let K be a chain complex composed of free groups. Let G be finitely generated. Then the sequence (*) of 1.2A is exact. This exact sequence is natural with respect to chain maps $f: K \to K'$ and homomorpisms $\phi: G \to H$.*

Proof. Let $0 \to R \xrightarrow{i} F \xrightarrow{j} G \to 0$ be exact, where F is a finitely generated free abelian group. R is finitely generated and free also. By 1.3A, the sequence corresponding to this exact coefficient sequence is exact:

$$\cdots \to H^r(K; R) \xrightarrow{i_*} H^r(K; F) \xrightarrow{j_*} H^r(K; G) \xrightarrow{\delta_*} H^{r+1}(K; R)$$
$$\xrightarrow{i_*} H^{r+1}(K; F) \to \cdots.$$

Hence the following sequence is exact:

$$0 \to \operatorname{Ker} \delta_* \to H^r(K;G) \to \operatorname{Im} \delta_* \to 0.$$

However, $\operatorname{Ker} \delta_* = \operatorname{Im} j_* \approx \operatorname{Coker} i_*$ and $\operatorname{Im} \delta_* = \operatorname{Ker} i_*$ by exactness. By 2.1A, we see that

$$\operatorname{Coker} i_* \approx \operatorname{Coker} (H^r(K) \otimes R \to H^r(K) \otimes F) \approx H^r(K) \otimes G$$

and

$$\operatorname{Ker} i_* \approx \operatorname{Ker} (H^{r+1}(K) \otimes R \to H^{r+1}(K) \otimes F) = \operatorname{Tor}(H^{r+1}(K), G)$$

(see [5; p. 160] for the definition of Tor). Under these isomorphisms, the inclusion $\operatorname{Ker} \delta_* \to H^r(K;G)$ goes over to $\alpha\colon H^r(K) \otimes G \to H^r(K;G)$, and δ_* defines $\beta\colon H^r(K;G) \to \operatorname{Tor}(H^{r+1}(K), G)$. Hence the sequence (*) is exact. The naturality statements follow from the naturality statements of 1.3A.

COROLLARY 2.3. *The universal coefficient theorem for cohomology*

$$0 \to H^r(X,A) \otimes G \xrightarrow{\alpha} H^r(X,A;G) \xrightarrow{\beta} \operatorname{Tor}(H^{r+1}(X,A), G) \to 0$$

holds in the following cases:

1) *simplicial (or cellular) theory for finite complexes and G arbitrary, or not necessarily finite complexes and G finitely generated:*

2) *singular theory for G finitely generated; and*

3) *Cech theory for (X,A) compact and G arbitrary, or (X,A) paracompact and G fiinitely generated.*

Proof. This follows immediately from 1.2A and 2.2A and the fact that direct limits preserve exactness, $\otimes$, and Tor [4].

3. A counter-example. This example shows that the exact sequence of 2.3A does not hold in general for singular theory or cellular theory. Let X be an $(n-1)$-connected CW-complex such that $H_n(X) = Q =$ the additive group of rationals. Let $G = Q$. By 1.1A, $H^n(X;Z) = \operatorname{Hom}(Q,Z) = 0$, and $H^n(X;Q) = \operatorname{Hom}(Q,Q) = Q$. However, 2.3A would give that

$$0 \to 0 \otimes Q \to Q \to \operatorname{Tor}(H^{n+1}(X), Q) \to 0$$

is exact; but $\operatorname{Tor}(H^{n+1}(X), Q) = 0$, and hence $Q = 0$. This is a contradiction.

PRINCETON UNIVERSITY.

REFERENCES.

[1] A. Borel and J.-P. Serre, "Groupes de Lie et Puissances Réduites de Steenrod," *American Journal of Mathematics*, vol. 75 (1953), pp. 409-448.

[2] K. Borsuk, "Sur les groupes des classes de transformations continues," *Comptes Rendus de l'Academie Sciences*, Paris, vol. 202 (1936), pp. 1400-1403.

[3] H. Cartan, "Sur les groupes d'Eilenberg-MacLane," I and II, *Proceedings of the National Academy of Sciences*, vol. 40 (1954), pp. 467-471, 704-707.

[4] H. Cartan and S. Eilenberg, *Homological Algebra*, Princeton University Press, 1956.

[5] S. Eilenberg and N. E. Steenrod, *Foundations of Algebraic Topology*, Princeton University Press, 1953.

[6] P. J. Hilton, *An Introduction to Homotopy Theory*, Cambridge University Press, 1953.

[7] H. Hopf, "Die klassen der abbildungen der n-dimensionalen Polyeder der n-dimensionale Sphäre," *Commentarii Mathematici Helvetici*, vol. 5 (1953), pp. 39-54.

[8] W. S. Massey, "Exact couples in algebraic topology," I and II, III, IV, and V, *Annals of Mathematics*, vol. 56 (1952), pp. 363-396, vol. 57 (1953), pp. 248-286.

[9] W. S. Massey, "Some problems in algebraic topology and the theory of fibre bundles," *Annals of Mathematics*, vol. 62 (1955), pp. 327-359.

[10] J.-P. Serre, "Homologie singulière des espaces fibrés," *Annals of Mathematics*, vol. 54 (1951), pp. 425-505.

[11] ———, "Groupes d'homotopie et classes de groupes Abéliens," *Annals of Mathematics*, vol. 58 (1953), pp. 258-294.

[12] ———, "Cohomologie modulo 2 des complexes d'Eilenberg-MacLane," *Commentarii Mathematici Helvetici*, vol. 27 (1953), pp. 198-323.

[13] E. H. Spanier, "Borsuk's cohomotopy groups," *Annals of Mathematics*, vol. 50 (1949), pp. 203-245.

[14] N. E. Steenrod, "Products of cocycles and extensions of mappings," *Annals of Mathematics*, vol. 48 (1947), pp. 290-320.

[15] ———, "Homology groups of symmetric groups and reduced power operations," and "Cyclic reduced powers of cohomology classes," *Proceedings of the National Academy of Sciences*, vol. 39 (1953), pp. 213-217, 217-223.

[16] G. W. Whitehead, "On the Freudenthal theorems," *Annals of Mathematics*, vol. 57 (1954), pp. 209-228.

GENERALIZED COHOMOTOPY GROUPS.*

By Franklin P. Peterson.[1]

1. Introduction. One of the fundamental problems of topology is the computation of $\pi(K;X)$, the set of homotopy classes of maps of a complex K into a space X. When K is an n-sphere S^n, then $\pi(K;X)=\pi_n(X)$, the familiar n-th homotopy group of X. When $X=S^n$ and the dimension of K is $\leqq 2n-2$, then $\pi(K;X)=\pi^n(K)$, the n-th cohomotopy group of K. The structure of $\pi^n(K)$ has been studied in [11].

In this paper, we shall introduce cohomotopy groups with coefficients in an abelian group G; namely, $\pi^n(K;G)=\pi(K;X)$, where the dimension of K is $\leqq 2n-2$ and X is a simply connected space whose homology is zero except that its n-th homology group is G. $\pi^n(K;G)$ is shown to be independent of the choice of such a space X. Notice that if $G=Z=$ the additive group of integers, then $\pi^n(K;Z)=\pi^n(K)$. For a fixed group G, the properties of $\pi^n(K;G)$ are analogous to those of $\pi^n(K)$, and the results of [11] are generalized so as to apply to $\pi^n(K;G)$. Furthermore, if G has no elements of order 2, then a homomorphism $\phi\colon G\to H$ induces a unique homomorphism $\phi_\#\colon \pi^n(K;G)\to\pi^n(K;H)$. Except for this restriction on G, $\phi_\#$ enjoys many of the same properties as the induced cohomology homomorphism $\phi_*\colon H^n(K;G)\to H^n(K;H)$. In particular, the sequence corresponding to an exact coefficient sequence is exact, and there is a universal coefficient theorem which asserts that

$$0\to\pi^n(K)\otimes G\xrightarrow{\alpha}\pi^n(K;G)\xrightarrow{\beta}\operatorname{Tor}(\pi^{n+1}(K),G)\to 0$$

is a split exact sequence. This theorem reduces the problem of computing $\pi^n(K;G)$ to that of computing $\pi^n(K)$. We conclude with a section on cohomotopy operations and with some remarks on homotopy groups with coefficients in G dual to our generalized cohomotopy groups.

In conclusion, I wish to express my warm appreciation to Professor N. E. Steenrod for his kind advice and encouragement. This paper is essentially Part II of a paper written under his direction and submitted as a disserta-

* Received September 21, 1955.

[1] The author was a predoctoral National Science Foundation Fellow during the preparation of this paper.

tion to Princeton University. I also wish to thank Professor J. C. Moore for suggesting the idea of general coefficients.

2. Homotopy classes of maps. In this section, we review some known results on the existence of a group structure on the set of homotopy classes of maps $a\colon (K,L) \to (X,x_0)$.

Let (K,L) be a CW-pair with dimension $K = N$ (K is an N-dimensional CW-complex [7] and L is a subcomplex). Let X be an arcwise connected space, and let $x_0 \varepsilon X$. We denote by $\pi(K,L;X,x_0)$ the set of homotopy classes of maps $a\colon (K,L) \to (X,x_0)$. A map $f\colon (K,L) \to (K',L')$ induces a function $f^{\#}\colon \pi(K',L';X,x_0) \to \pi(K,L;X,x_0)$ defined by $f^{\#}([a]) = [af]$, where $[a]$ denotes the homotopy class of a. Also, a map $\phi\colon (X,x_0) \to X',x_0')$ induces a function $\phi_{\#}\colon \pi(K,L;X,x_0) \to \pi(K,L;X',x_0')$ defined by $\phi_{\#}([a]) = [\phi a]$.

If $X \neq \emptyset$, let SX denote the reduced suspension of X [15; p. 656]; namely, SX is the space obtained from $X \times I$ by identifying $X \times \{0\} \cup X \times \{1\} \cup \{x_0\} \times I$ to a point x_0. (x_0 is used to denote the base point of both X and SX.) Let $\mathcal{S}X$ denote the suspension of X [16]; namely, $\mathcal{S}X$ is the space obtained from $X \times I$ by identifying $X \times \{0\}$ and $X \times \{1\}$ to points. If $X = \emptyset$, define $S\emptyset$ and $\mathcal{S}\emptyset$ to be a pair of points. Also define $S^rX = S(S^{r-1}X)$, $S^0X = X$, $\mathcal{S}^rX = \mathcal{S}(\mathcal{S}^{r-1}X)$, and $\mathcal{S}^0X = X$. Let $\mu\colon (\mathcal{S}K, \mathcal{S}L) \to (SK, SL)$ be the canonical map contracting $\{x_0\} \times I$ to x_0. Since (K,L) is a CW-pair, μ is a homotopy equivalence and thus

$$\mu^{\#}\colon \pi(SK,SL;X,x_0) \to \pi(\mathcal{S}K, \mathcal{S}L;X,x_0)$$

is a 1-1 correspondence.

As in [16], suspension induces a function

$$S_{\#}\colon \pi(K,L;X,x_0) \to \pi(SK,SL;SX,x_0).$$

$S_{\#}$ is natural with respect to maps $f\colon (K,L) \to (K',L')$ and maps $\phi\colon (X,x_0) \to (X',x_0')$.

THEOREM 2.1. *If X is an $(n-1)$-connected space and (K,L) is a CW-pair with dimension $K \leqq 2n-2$, then $S_{\#}$ is a 1-1 correspondence.*

Proof. This is an immediate consequence of Corollary 7.2 of [16].

Let $(X,x_0)^{K,L}$ denote the function space of maps $a\colon (K,L) \to (X,x_0)$ with the compact-open topology. The constant map at x_0 serves as the base point for this function space. In [1], a function

$$\lambda\colon \pi_r((X,x_0)^{K,L}) \to \pi(S^rK, S^rL;X,x_0)$$

is defined and the following theorem is proven [1; p. 81]:

THEOREM 2.2. *λ is a 1-1 correspondence. This correspondence is natural with respect to maps $f\colon (K,L)\to(K',L')$ and maps $\phi\colon (X,x_0)\to(X',x_0')$.*

As in [16], we now prove the main result on the existence of a group structure on $\pi(K,L;X,x_0)$.

THEOREM 2.3. *If X is an $(n-1)$-connected space and (K,L) is a CW-pair with dimension $K\leqq 2n-2$, then $\pi(K,L;X,x_0)$ is an abelian group. This group structure is natural with respect to maps $f\colon (K,L)\to(K',L')$ and maps $\phi\colon (X,x_0)\to(X',x_0')$.*[2]

Proof. By 2.1, $S_\#^2\colon \pi(K,L;X,x_0)\to\pi(S^2K,S^2L;S^2X,x_0)$ is a natural 1-1 correspondence. Also, by 2.2,

$$\lambda\colon \pi_2((S^2X,x_0)^{K,L})\to\pi(S^2K,S^2L;S^2X,x_0)$$

is a natural 1-1 correspondence, and $\pi_2((S^2X,x_0)^{K,L})$ is an abelian group. We define the group structure on $\pi(K,L;X,x_0)$ using the 1-1 correspondence $\lambda^{-1}S_\#^2$. (This addition of homotopy classes is analogous to that defined in [14].) Note that $S_\#$ is now an isomorphism. This group structure is natural with respect to maps $f\colon (K,L)\to(K',L')$ and maps $\phi\colon (X,x_0)\to(X',x_0')$ because the 1-1 correspondences λ and $S_\#$ are; i.e. these maps induce homomorphisms $f^\#$ and $\phi_\#$.

We now define the sequence of a CW-pair (K,L). Define

$$\Delta\colon \pi(L;X,x_0)\to\pi(K,L;SX,x_0)$$

as follows: let $a\in[a]\in\pi(L;X,x_0)$. Extend a to a map $a'\colon (K,L)\to(CX,X)$, where CX denotes the cone on X [15; p. 656]. Let $h\colon (CX,X)\to(SX,x_0)$ be the canonical map collapsing X to x_0 [15; p. 657]. The composition $ha'\colon (K,L)\to(SX,x_0)$ represents $\Delta([a])$. When (K,L) and (X,x_0) satisfy the conditions of Theorem 2.3, Δ is a natural homomorphism (Δ is strictly analogous to the homomorphism Δ for ordinary cohomotopy [14; p. 216]). Let $i\colon L\to K$ and $j\colon K\to(K,L)$ be inclusions. Then the cohomotopy sequence of the pair (K,L) is defined to be the following sequence of groups and homomorphisms:

$$\pi(K,L;X,x_0)\xrightarrow{j^\#}\pi(K;X,x_0)\xrightarrow{i^\#}\pi(L;X,x_0)\xrightarrow{\Delta}\pi(K,L;SX,x_0)$$
$$\xrightarrow{j^\#}\pi(K;SX,x_0)\to\cdots.$$

[2] When we say a structure is natural with respect to maps $f\colon (K,L)\to(K',L')$ and maps $\phi\colon (X,x_0)\to(X',x_0')$, we assume (K',L') and (X',x_0') satisfy the same dimensional or connectedness assumptions that (K,L) and (X,x_0) satisfy.

In Section 4 we prove that this sequence is exact when (K, L) and (X, x_0) satisfy the conditions of Theorem 2.3.

3. Cohomotopy groups with coefficients in G. In this section we define our generalized cohomotopy groups and state their elementary properties. In making the definition certain arbitrary choices are necessary, and we prove independence of these choices under certain restrictions on G. We also state our main results on generalized cohomotopy groups; the proofs are given later in the paper.

Let G be an abelian group and let $n > 1$. An $(n+1)$-dimensional CW-complex X is said to be an *$X(G, n)$-space* if $\pi_1(X) = 0$, $H_i(X) = 0$ for $i \neq n$, and $H_n(X) = G$. (This concept was introduced by Moore [10; p. 550].) Note that if X is an $X(G, n)$-space, then SX is an $X(G, n+1)$-space.

In Section 5 we prove

LEMMA 3.1. *For given G and n, there exists an $X(G, n)$-space.*

Let X be an $X(G, n)$-space, and let (K, L) be a CW-pair with dimension $K = N$. We define the *n-th cohomotopy group of (K, L) with coefficients in G* to be $\pi(K, L; X, x_0)$, and denote it by $\pi^n(K, L; G)$. When we use the notation $\pi^n(K, L; G)$, we assume $n > (N+1)/2$ and thus $\pi^n(K, L; G)$ has a natural group structure. As defined, $\pi^n(K, L; G)$ depends on the choice of $X(G, n)$-space. We show below that $\pi^n(K, L; G)$ is naturally independent of this choice when G has no elements of order 2. However, if we do not change coefficients during a discussion, it suffices to choose a fixed $X(G, t)$-space Y and use $S^{n-t}Y$ as the $X(G, n)$-space for each $n \geqq t \geqq 2$.

In Section 5 we prove

THEOREM 3.2. *$\pi^n(K, L; G)$ satisfies all the axioms for cohomology of Eilenberg and Steenrod* [6; p. 13] *in those dimensions where a natural group structure is defined.*

We now return to the question of independence of the choice of the $X(G, n)$-space. Let X be an $X(G, n)$-space, Y an $X(H, n)$-space. There is a natural homomorphism $\eta\colon \pi(X, x_0; Y, y_0) \to \operatorname{Hom}(G, H)$ defined by $\eta([a]) = a_*\colon H_n(X, x_0) = G \to H_n(Y, y_0) = H$. In Section 5 we prove

THEOREM 3.3. *If $n \geqq 3$, then $\eta\colon \pi(X, x_0; Y, y_0) \to \operatorname{Hom}(G, H)$ is an epimorphism and has a kernel isomorphic to* $\operatorname{Ext}(G, H \otimes Z_2)$.[3]

Let $\mathfrak{D}$ be the family of abelian groups having no elements of order 2. In the appendix we prove

[3] See [5] or [6] for the definition and properties of Hom and Ext.

LEMMA 3.4. *If* $G \varepsilon \mathfrak{D}$, *then* $\operatorname{Ext}(G, H \otimes Z_2) = 0$.

Let $\phi\colon G \to H$ be a homomorphism, and let X be an $X(G,n)$-space, Y an $X(H,n)$-space. There exists a map ${}^*\phi\colon (X, x_0) \to (Y, y_0)$ such that $({}^*\phi)_* = \phi\colon H_n(X, x_0) \to H_n(Y, y_0)$ by 3.3. In fact, if $G \varepsilon \mathfrak{D}$, then ${}^*\phi$ is unique up to homotopy by 3.3 and 3.4. Furthermore, if $\psi\colon H \to J$, Z is an $X(J,n)$-space, and $H \varepsilon \mathfrak{D}$ also, then ${}^*\psi{}^*\phi \simeq {}^*(\psi\phi)\colon (X, x_0) \to (Z, z_0)$.

Now let X and X' be two different $X(G,n)$-spaces, let $G \varepsilon \mathfrak{D}$, and let $\phi\colon G \to G$ be the identity homomorphism. Then there exist maps ${}^*\phi\colon (X, x_0) \to (X', x_0')$ and ${}^*\phi'\colon (X', x_0') \to (X, x_0)$ inducing ϕ such that $({}^*\phi')({}^*\phi)$ and $({}^*\phi)({}^*\phi')$ are homotopic to the identity maps. Furthermore, ${}^*\phi$ and ${}^*\phi'$ are unique up to homotopy. Hence ${}^*\phi$ and ${}^*\phi'$ induce unique isomorphisms

$$({}^*\phi)_\#\colon \pi(K, L; X, x_0) \to \pi(K, L; X', x_0')$$

and

$$({}^*\phi')_\#\colon \pi(K, L; X', x_0') \to \pi(K, L; X, x_0)$$

which are inverses of each other. Hence the set of groups $\{\pi(K, L; X, x_0)\}$ for all $X(G,n)$-spaces X form a transitive system of groups [6; p. 17], and we have shown that $\pi^n(K, L; G)$ is independent of the choice of $X(G,n)$-space. In case $G \not\varepsilon \mathfrak{D}$, $({}^*\phi)_\#$ is an isomorphism, but it is not a unique isomorphism. In this case, we assume a fixed $X(G,2)$-space during any given discussion.

As a further corollary of the above discussion, if $G \varepsilon \mathfrak{D}$, then a homomorphism $\phi\colon G \to H$ induces a unique homomorphism $\phi_\#\colon \pi^n(K, L; G) \to \pi^n(K, L; H)$. This is natural in the sense that if $\phi\colon G \to H$, $\psi\colon H \to J$ and G and $H \varepsilon \mathfrak{D}$, then $(\psi\phi)_\# = \psi_\#\phi_\#$.

We now state our main results on cohomotopy groups with coefficients in G. Let G' and $G \varepsilon \mathfrak{D}$, and let $0 \to G' \xrightarrow{\phi} G \xrightarrow{\psi} G'' \to 0$ be an exact coefficient sequence. In Section 6 we define a homomorphism

$$\delta_\#\colon \pi^r(K, L; G'') \to \pi^{r+1}(K, L; G')$$

for $r > (N+1)/2$. Define the sequence corresponding to this exact coefficient sequence to be the following sequence of groups and homomorphisms:

$$\cdots \to \pi^r(K, L; G') \xrightarrow{\phi_\#} \pi^r(K, L; G) \xrightarrow{\psi_\#} \pi^r(K, L; G'') \xrightarrow{\delta_\#} \pi^{r+1}(K, L; G') \to \cdots.$$

In Section 6 we prove

THEOREM 3.5. *For $r > (N+1)/2$, the sequence corresponding to an exact coefficient sequence is exact. This sequence is natural with respect to maps $f\colon (K,L) \to (K',L')$, and if G', G, G'', H', and $H \in \mathfrak{D}$, then it is natural with respect to a homomorphism of one exact sequence into another:*

$$\begin{array}{ccccccccc} 0 \to & G' & \xrightarrow{\phi} & G & \xrightarrow{\psi} & G'' & \to 0 \\ & \downarrow & & \downarrow & & \downarrow & \\ 0 \to & H' & \xrightarrow{\phi'} & H & \xrightarrow{\psi'} & H'' & \to 0. \end{array}$$

The above is a generalization of an exact sequence of Moore [10; p. 552].

Define a function $\pi^r(K,L) \times G \to \pi^r(K,L;G)$ by $([a],[g]) \to [ga]$, where $[g] \in \pi_r(X,x_0) = G$, $[a] \in \pi^r(K,L)$, and X is an $X(G,r)$-space. By 2.3, this is a bilinear function for $r > (N+1)/2$, and hence it induces a homomorphism $\alpha\colon \pi^r(K,L) \otimes G \to \pi^r(K,L;G)$. In Section 7 we define a homomorphism $\beta\colon \pi^r(K,L;G) \to \operatorname{Tor}(\pi^{r+1}(K,L),G)$ and prove the following universal coefficient theorems:

THEOREM 3.6. *Let G be a finitely generated abelian group. Then the sequence*

$$(*) \qquad 0 \to \pi^r(K,L) \otimes G \xrightarrow{\alpha} \pi^r(K,L;G) \xrightarrow{\beta} \operatorname{Tor}(\pi^{r+1}(K,L),G) \to 0$$

is exact for $r > (N+1)/2$. This sequence is natural with respect to maps $f\colon (K,L) \to (K',L')$, and if $G \in \mathfrak{D}$, then it is natural with respect to homomorphisms $\phi\colon G \to H$. Furthermore, if $\pi^r(K,L)$ is finitely generated and $G \in \mathfrak{D}$, then the exact sequence $(*)$ splits.

THEOREM 3.7. *Let $G \in \mathfrak{D}$, and let (K,L) be a finite CW-pair. Then the sequence $(*)$ is exact for $r > (N+1)/2$. This sequence is natural with respect to maps $f\colon (K,L) \to (K',L')$ and homomorphisms $\phi\colon G \to H$.*

In Section 7 we give an example to show that if G is not finitely generated and (K,L) is not finite, then the sequence $(*)$ is not necessarily exact.

Theorem 3.6 generalizes an exact sequence of Serre [13; p. 284]; Serre's sequence is 3.6 for the case $(K,L) = (S^n, \text{pt.})$ and $G =$ a cyclic group. He notes that his sequence does not split when $G = Z_2$. It is a result of Barratt [2; p. 283] that Serre's sequence splits for G a cyclic group other than Z_2.

Theorems 3.6 and 3.7, beside giving a further analogy between cohomotopy groups with arbitrary coefficients and cohomology groups with arbitrary coefficients, reduce the problem of calculating $\pi^r(K,L;G)$ to the standard

problem of calculating $\pi^r(K,L)$. Hence, Theorems 3.6 and 3.7 are two of our main results.

Let X be an $X(G,n)$-space. $S_\#: \pi_{n+s}(X) \to \pi_{n+s+1}(SX)$ is an isomorphism when $s < n-1$ by Theorem 2.1. We identify these groups under this isomorphism and denote the result by $G_{(s)}$. For any class $\mathscr{C}$ [13], let $\alpha(\mathscr{C}; G)$ denote the largest integer such that $G_{(s)} \in \mathscr{C}$ for $0 < s < \alpha(\mathscr{C}; G)$ (see definition in [11]). In Section 4 we define a natural homomorphism $\eta^r: \pi^r(K,L;G) \to H^r(K,L;G)$ analogous to η^r for ordinary cohomotopy. In Section 8 we prove the following generalization of Theorem 3.1 of [11]:

THEOREM 3.8. *Let $\mathscr{C}$ be a class satisfying condition* (II_B) (*see* [13] *or* [11]), *and let G be finitely generated. Let $n > (N+1)/2$ be such that $H^r(K,L;G) \in \mathscr{C}$ for every $r > n$. Then $\eta^r: \pi^r(K,L;G) \to H^r(K,L;G)$ is a $\mathscr{C}$-isomorphism if $r > \mathrm{Max}((N+1)/2, n-\alpha(\mathscr{C};G))$, and is a $\mathscr{C}$-epimorphism for $r = n - \alpha(\mathscr{C};G)$ in case $n-\alpha(\mathscr{C};G) > (N+1)/2$.*

Theorems 2.1 of [11] and 3.6 give information on $\alpha(\mathscr{C};G)$ for various $\mathscr{C}$ and G, and we may draw consequences of 3.8 similar to 3.2, 3.3, 3.4, and 3.5 of [11]. Also, the result of Adem carries over to general coefficients and there is a theorem analogous to Theorem 3.8 of [11]. Furthermore, for the case $G = Z_p$, there is a theorem analogous to Theorem 3.9 of [11]. Rather than considering these in detail, let us note that any result on the structure of $\pi^r(K,L)$ gives a result on the structure of $\pi^r(K,L;G)$ by 3.6 or 3.7.

4. The exact sequences of a pair. In this section we prove the exactness of the cohomotopy sequence of a pair defined in Section 2. We then show that the cohomotopy exact couple of [11] can be generalized; some of our main results are based on this generalized cohomotopy exact couple.

In the cohomotopy sequence of the pair (K,L), the following relations are obvious: $\mathrm{Im}\, j^\# = \mathrm{Ker}\, i^\#$, $\mathrm{Im}\, i^\# \subset \mathrm{Ker}\, \Delta$, and $\mathrm{Im}\, \Delta \subset \mathrm{Ker}\, j^\#$. To prove $\mathrm{Im}\, i^\# = \mathrm{Ker}\, \Delta$ and $\mathrm{Im}\, \Delta = \mathrm{Ker}\, j^\#$, we need the assumptions that X is $(n-1)$-connected and dimension $K = N \leqq 2n-2$. The exactness now follows from Theorem $(3.1)_0$ of [15; p. 657] with the carrier ϕ from K to $S^{N-n+1}X$ being such that $\phi L = x_0$ and with the unrestricted carrier ψ from K to $S^{N-n+1}X$. The complete details are left to the reader. Note that the exact sequence of a pair (K,L) gives immediately the exact cohomotopy sequence of a triple (K,L,M) [6; p. 25].

The results of Section 4 of [11] on the cohomotopy exact couple now carry over to our more general situation. Let X be an $(n-1)$-connected space. We replace S^{n+t} in [11] by S^tX for $t \geqq 0$. With this substitution,

all of the results of Section 4 of [11] are true; the main identifications now being that $\mathcal{H}^{r,s} \approx H^s(K, L; \pi_{s-r+n}(X))$ for $r > \mathrm{Max}(n, (N+1)/2)$ and $\Gamma^{r,r-1} \approx \pi(K, L; S^{r-n}X, x_0)$ for $r > \mathrm{Max}(n-1, (N+1)/2)$. Again we denote $i'^{r,r}\colon \pi(K, L; S^{r-n}X, x_0) \to H^r(K, L; \pi_n(X))$ by η^r.

For later use, we make the following definition. Let X be an $(n-1)$-connected space. Let dimension $K = N \leqq 2(s-r+n) - 2$ and let $s \geqq r$. We define

$$\theta\colon \pi_r((S^sX, x_0)^{K,L}) \to \pi(K, L; S^{s-r}X, x_0)$$

to be the following composition:

$$\pi_r((S^sX, x_0)^{K,L}) \xrightarrow{\lambda} \pi(\mathcal{S}^rK, \mathcal{S}^rL; S^sX, x_0) \xrightarrow{\mu^{\#-1}} \pi(S^rK, S^rL; S^sX, x_0)$$
$$\xrightarrow{(S_{\#}^{-1})^r} \pi(K, L; S^{s-r}X, x_0).$$

λ is the isomorphism of 2.2, $S_{\#}^{-1}$ is the isomorphism of 2.1 ($S_{\#}$ is an isomorphism under the above restrictions on s, r, n, and N), and $\mu^{\#}$ is the isomorphism of Section 2. θ is natural with respect to maps $f\colon (K, L) \to (K', L')$ and maps $\phi\colon (X, x_0) \to (X', x_0')$ because λ, $\mu^{\#}$, and $S_{\#}$ are.

5. Independence of the choice of $X(G, n)$-space. In this section we give proofs of 3.1, 3.2, and 3.3 as well as giving some further properties of $\pi^n(K, L; G)$. These further properties will be used in Section 7.

Proof of 3.1. Let $G = F/R$, where F is a free abelian group on generators $\{x_\alpha\}_{\alpha \varepsilon A}$ and R is the group of relations. R is free abelian [6; p. 134] with basis $\{y_\beta\}_{\beta \varepsilon B}$. Let T be the CW-complex $\bigvee_{\alpha \varepsilon A} S^n_\alpha$, a union of n-spheres S^n_α with a single point in common. By the Hurewicz theorem, $\pi_n(T) \overset{\eta}{\approx} H_n(T) = F$. For each $\beta \varepsilon B$, attach an $(n+1)$-cell e_β to T by a map representing $\eta^{-1}(y_\beta) \varepsilon \pi_n(T)$. Call the resulting space X. By construction, $\pi_1(X) = 0$, $H_i(X) = 0$ for $i < n$, $\pi_n(X) \approx H_n(X) = G$, and $H_i(X) = 0$ for $i > n+1$. Moreover, since any non-zero $(n+1)$-cycle in X would imply a non-trivial relation among the $\{y_\beta\}_{\beta \varepsilon B}$, $H_{n+1}(X) = 0$.

In order to prove 3.3, we first prove

LEMMA 5.1. *Let X be an $X(G, n)$-space. Then $\pi_{n+1}(X) \approx G \otimes Z_2$.*

Proof. The proof is based on figure 1, a portion of the first derived homotopy exact couple of Massey [8; part II]. $H_{n+1}(X) = 0$ and hence by

exactness $\pi_{n+1}(X, x_0) \approx \Gamma \approx H_n(X; Z_2)$. By this universal coefficient theorem, $H_n(X; Z_2) \approx G \otimes Z_2$.

$$
\begin{tikzcd}
 & & \vdots \arrow[d] & & & \\
 & \cdots \arrow[r] & 0 \arrow[r] \arrow[d] & \cdots & \vdots \arrow[d] & \\
\cdots \arrow[r] & 0 \arrow[r] & \Gamma \arrow[r] \arrow[d] & H_n(X; Z_2) \arrow[r] & 0 \arrow[r] \arrow[d] & \cdots \\
 & \cdots \arrow[r] & \pi_{n+1}(X, x_0) \arrow[r] \arrow[d] & H_{n+1}(X) \arrow[r] & 0 \arrow[r] \arrow[d] & \cdots \\
 & & \vdots & & \vdots &
\end{tikzcd}
$$

Figure 1.

Proof of 3.3. The proof is based on figure 2, a portion of the first derived generalized cohomotopy exact couple. Let $(K, L) = (X, x_0)$, $N = n + 1$, and let (Y, y_0) be the coefficient space. Now

$$H^n(X, x_0; \pi_n(Y, y_0)) \approx H^n(X, x_0; H) \approx \mathrm{Hom}(G, H)$$

by the universal coefficient theorem. By 5.1 and the universal coefficient theorem,

$$H^{n+1}(X, x_0; \pi_{n+1}(Y, y_0)) \approx H^{n+1}(X, x_0; H \otimes Z_2) \approx \mathrm{Ext}(G, H \otimes Z_2).$$

By exactness η is an epimorphism with kernel isomorphic to

$$\Gamma^{n,n} \approx \mathrm{Ext}(G, H \otimes Z_2).$$

$$
\begin{tikzcd}
 & & \vdots \arrow[d] & & & \\
 & \cdots \arrow[r] & 0 \arrow[r] \arrow[d] & \cdots & \vdots \arrow[d] & \\
\cdots \arrow[r] & 0 \arrow[r] & \Gamma^{n,n} \arrow[r, "i'"] \arrow[d, "j'"] & H^{n+1}(X, x_0; \pi_{n+1}(Y, y_0)) \arrow[r] & 0 \arrow[r] \arrow[d] & \cdots \\
 & \cdots \arrow[r] & \pi(X, x_0; Y, y_0) \arrow[r, "\eta"] \arrow[d] & H^n(X, x_0; \pi_n(Y, y_0)) \arrow[r] & 0 \arrow[r] \arrow[d] & \cdots \\
 & & \vdots & & \vdots &
\end{tikzcd}
$$

Figure 2.

Before we prove 3.2, let us draw some further corollaries of 3.3.

Corollary 5.2. *If $G \in \mathfrak{D}$, then $\pi^n(K, L; G)$ is a unitary left module over any subring of the ring of endomorphisms of G.*

Proof. By the naturality statements of Theorem 2.3, $\pi^n(K,L;G)$ is a unitary left module over $\pi(X,x_0;X,x_0)$ where X is an $X(G,n)$-space (the multiplication is composition). By 3.3 and 3.4, $\eta\colon \pi(X,x_0;X,x_0) \to \mathrm{Hom}(G,G)$ is an isomorphism if $G \in \mathfrak{D}$, and the corollary follows.

COROLLARY 5.3. *If G is a field and $G \in \mathfrak{D}$, then $\pi^n(K,L;G)$ is a vector space over G.*

Proof. Any field is a subring of its ring of additive endomorphisms; namely, $g \in G$ acts on G by left multiplication. The corollary now follows from Corollary 5.2.

We now complete the proof of 3.2. The results of Sections 2 and 4 show that $\pi^n(K,L;G)$ satisfy axioms 1 through 4 of Eilenberg and Steenrod.

Let (K,L) be a CW-pair with $L \neq \emptyset$. Let (K_L, p_L) be the CW-pair obtained by identifying L to a point p_L. Let $f\colon (K,L) \to (K_L, p_L)$ be the canonical map.

LEMMA 5.4. *$f^{\#}\colon \pi(K_L,p_L;X,x_0) \to \pi(K,L;X,x_0)$ is a 1-1 correspondence.*

Proof. Same as in [14; p. 215].

Let (K,L) be a CW-pair. Let $M \subset L$ be such that $K - M$ is a subcomplex of K. Let $i\colon (K-M, L-M) \to (K,L)$ be the inclusion.

THEOREM 5.5. (Excision Axiom).

$$i^{\#}\colon \pi^r(K,L;G) \to \pi^r(K-M, L-M;G)$$

is an isomorphism for $r > (N+1)/2$.

Proof. This follows immediately from 5.4 as in [14; p. 215].

Let $f,g\colon (K,L) \to (K',L')$ be homotopic.

THEOREM 5.6. (Homotopy Axiom).

$$f^{\#} \text{ and } g^{\#}\colon \pi^r(K',L';G) \to \pi^r(K,L;G)$$

are equal.

Proof. Obvious.

Theorem 3.2 has now been proven. A corollary to this theorem is the fact that any theorem derivable from the axioms for cohomology of Eilenberg and Steenrod which does not make use of the lower dimensional groups holds for cohomotopy groups with coefficients in G. An example of this is the Mayer-Vietoris sequence of a triad [6; p. 39].

6. Proof of Theorem 3. 5. This section is devoted to the proof that the sequence corresponding to an exact coefficient sequence is exact. The preliminary results are only steps in this proof.

LEMMA 6.1. *Let $G \varepsilon \mathfrak{D}$, and let $n > (N+1)/2$. Let X be an $X(G,n)$-space. Let Y be an $(n-1)$-connected space such that $H_n(Y,y_0) = G$ and $H_r(Y,y_0) = 0$ for $n < r < 2n$. Then there is a map $k\colon (X,x_0) \to (Y,y_0)$, unique up to homotopy, such that $k_*\colon H_n(X,x_0) = G \to H_n(Y,y_0) = G$ is the identity and $k_\#\colon \pi^n(K,L;G) \to \pi(K,L;Y,y_0)$ is a natural isomorphism.*

Proof. Similarly to the argument used in the proof of 3.3, there is a map $k\colon (X,x_0) \to (Y,y_0)$ such that $k_*\colon H_n(X,x_0) = G \to H_n(Y,y_0) = G$ is the identity. Hence $k_*\colon H_r(X,x_0) \to H_r(Y,y_0)$ is an isomorphism for $r < 2n$; it follows by a theorem of J. H. C. Whitehead [18; p. 215] that $k_\#\colon \pi_r(X,x_0) \to \pi_r(Y,y_0)$ is an isomorphism for $r < 2n-1$ and an epimorphism for $r = 2n-1$. Assuming that k is an inclusion by the mapping cylinder construction, this means that the pair (Y,X) is $(2n-1)$-connected [3; p. 183]. The lemma is now as easy consequence of the deformation obstruction theory as described in [3; p. 186]. The details of the proof are left to the reader.

Let $p\colon E \to B$ be a fibre space in the sense of Serre, with fibre F over $b_0 \varepsilon B$. Let K be a CW-complex. Define a map $p_1\colon E^K \to B^K$ by $p_1(a) = pa$.

LEMMA 6.2. *$p_1\colon E^K \to B^K$ is a fibre space with fibre F^K. A map $f\colon K \to K'$ induces a fibre preserving map ${}^*f\colon (E^{K'}, p_1', B^{K'}) \to (E^K, p_1, B^K)$, and a fibre preserving map $\phi\colon (E,p,B) \to (E',p',B')$ induces a fibre preserving map ${}^*\phi\colon (E^K, p_1, B^K) \to (E'^K, p_1', B'^K)$.*

Proof. The proof is a straight forward application of the covering homotopy theorem for (E,p,B), and the details are omitted.

Let G' and $G \varepsilon \mathfrak{D}$, and let $0 \to G' \xrightarrow{\phi} G \xrightarrow{\psi} G'' \to 0$ be an exact coefficient sequence. Let X be an $X(G, N+1)$-space, and let X'' be an $X(G'', N+1)$-space. Let ${}^*\psi\colon (X,x_0) \to (X'',x_0'')$ be a map inducing ψ. We replace ${}^*\psi$ by a fibre mapping as follows: assume ${}^*\psi$ is an inclusion by the mapping cylinder construction. Let Y be the space of paths in X'' which end in X. X is contained in Y as a deformation retract [6; p. 30] by $x \to$ constant path at x. Define $p\colon Y \to X''$ by $p(f) = f(0)$; p is a fibre map [12; p. 479] with fibre $F' =$ the space of paths starting at $x_0 \varepsilon X''$ and ending in X. We may assume $x_0 \varepsilon X \subset X''$. p is our replacement of ${}^*\psi$.

Using the technique of spectral sequences as applied to fibre spaces by Serre, we now prove the following lemma.

LEMMA 6.3. *$H_{N+1}(F') = G'$, $H_r(F') = 0$ otherwise for $r < 2(N+1) - 1$.*

Proof. Let E be the space of paths in X'' starting at x_0; E is a contractible space. Let $p_1\colon (E, F') \to (X'', X)$ be defined by $p_1(f) = f(1)$. This is a relative fibre space with fibre $F =$ the space of loops in X'' at x_0. F is $(N-1)$-connected because X'' is N-connected. As in [9; p. 330], there is a spectral sequence of this relative fibre space with $E_2^{p,q} \approx H_p(X'', X; H_q(F))$ and E_∞ is the graded group associated with $H(E, F')$. From the exact homology sequence of the pair (X'', X):

$$\cdots \to H_{N+2}(X'') \to H_{N+2}(X'', X) \to H_{N+1}(X) \to H_{N+1}(X'') \to H_{N+1}(X'', X) \to H_N(X) \to \cdots$$

we see that $H_{N+2}(X'', X) = G'$ and $H_r(X'', X) = 0$ otherwise. Hence $E_2^{p,q} = 0$ for $q < N$ except that $E_2^{N+2,0} = G'$, and $E_2^{p,q} = 0$ for $p < N+2$. Thus $H_r(E, F') = 0$ for $r < 2N + 2$ except that $H_{N+2}(E, F') = G'$. However E is contractible, and hence by the exact homology sequence of the pair (E, F'), $H_{N+1}(F') = G'$ and $H_r(F') = 0$ otherwise for $r < 2N + 1$.

In order to prove the naturality of Theorem 3.5, we first prove the following lemma.

LEMMA 6.4. *Given a homomorphism of one exact coefficient sequence into another:*

$$\begin{tikzcd} 0 \arrow[r] & G' \arrow[r, "\phi"] \arrow[d, "\xi'"] & G \arrow[r, "\psi"] \arrow[d, "\xi"] & G'' \arrow[r] \arrow[d, "\xi''"] & 0 \\ 0 \arrow[r] & H' \arrow[r, "\phi'"] & H \arrow[r, "\psi'"] & H'' \arrow[r] & 0. \end{tikzcd}$$

Let G', G, G'', H', and $H \in \mathfrak{D}$. Refer to the above construction. Then there is a fibre preserving map $\xi\colon (Y, p, X'') \to (Y_H, p, X''_H)$ which is homotopic to $^*\xi$ on Y and to $^*\xi''$ on X''.

Proof.

$$\begin{tikzcd} G \arrow[r, "\psi"] \arrow[d, "\xi"] & G'' \arrow[d, "\xi''"] \\ H \arrow[r, "\psi'"] & H'' \end{tikzcd} \quad \text{is commutative and hence} \quad \begin{tikzcd} X \arrow[r, "^*\psi"] \arrow[d, "^*\xi"] & X'' \arrow[d, "^*\xi''"] \\ X_H \arrow[r, "^*\psi'"] & X''_H \end{tikzcd}$$

is commutative up to homotopy by the results of Section 3. Assume $^*\psi$ and $^*\psi'$ are inclusions. $^*\xi\colon X \to X_H$ is such that $(^*\psi')(^*\xi) \simeq (^*\xi'')(^*\psi)$, i.e. $^*\xi$ is homotopic in X''_H to a map which can be extended to all of X'', namely $^*\xi''|X$. Use the homotopy extension theorem to define a map $\xi_1\colon (X'', X)$

$\to (X''_H, X_H)$ such that $\xi_1 | X'' \simeq {}^*\xi''$ and $\xi_1 | X \simeq {}^*\xi$. ξ_1 induces a fibre preserving map $\xi\colon (Y, p, X'') \to (Y_H, p, X''_H)$ having the correct properties. This completes the proof.

We now prove Theorem 3.5.

Proof of 3.5. By 5.4 and the exact cohomotopy sequence of the pair (K_L, p_L), it suffices to prove this theorem for the case $L = \emptyset$. Use the above construction and 6.2 to obtain the fibre space (Y^K, p_1, X''^K) with fibre F'^K. Consider the homotopy sequence of this fibre space and use the isomorphism θ of Section 4:

$$\begin{array}{ccccccccc}
\cdots \to \pi_{N+1-r}(F'^K) & \xrightarrow{j'_\#} & \pi_{N+1-r}(Y^K) & \xrightarrow{p_{1\#}} & \pi_{N+1-r}(X''^K) & \xrightarrow{\partial} & \pi_{N-r}(F'^K) \to \cdots \\
\downarrow k_\#^{-1}\theta & & \downarrow \theta & & \downarrow \theta & & \downarrow k_\#^{-1}\theta \\
\cdots \to \pi^r(K; G') & \xrightarrow{\phi_\#} & \pi^r(K; G) & \xrightarrow{\psi_\#} & \pi^r(K; G) & \xrightarrow{\delta_\#} & \pi^{r+1}(K; G') \to \cdots
\end{array}$$

∂ is the boundary operator in the homotopy sequence of the fibre space, $j'_\#$ is induced by the inclusion $j'\colon F'^K \to Y^K$, and $p_{1\#}$ is induced by p_1. θ is defined and is an isomorphism for $r > (N+1)/2$ because

$$2(N+1-N-1+r)-2 = 2r-2 > 2(N+1)/2-2 = N-1.$$

$k_\#\colon \pi^r(K; G') \to \pi(K; F')$ is the isomorphism of 6.1 and 6.3, and $\delta_\#\colon \pi^r(K; G'') \to \pi^{r+1}(K; G')$ is defined by $\delta_\# = k_\#^{-1}\theta\partial\theta^{-1}$. Furthermore, $\theta j'_\# = \phi_\# k_\#^{-1}\theta$ and $\psi_\#\theta = \theta p_{1\#}$ by the naturality of θ. Also, $\delta_\#\theta = k_\#^{-1}\theta\delta$ by definition. Hence the sequence corresponding to the exact coefficient sequence is exact.

A map $f\colon K \to K'$ induces a fibre preserving map

$$f\colon (Y^{K'}, p_1', X''^{K'}) \to (Y^K, p_1, X''^K)$$

and hence $f^\#$ commutes with $\delta_\#$ because ∂ and θ are natural. $f^\#$ commutes with $\phi_\#$ and $\psi_\#$ also. Thus f induces a homomorphism of the sequence of K' into that of K. Also, a homomorphism of one exact coefficient sequence into another

$$\begin{array}{ccccccccc}
0 \to & G' & \xrightarrow{\phi} & G & \xrightarrow{\psi} & G'' & \to 0 \\
& \downarrow \xi' & & \downarrow \xi & & \downarrow \xi'' & \\
0 \to & H' & \xrightarrow{\phi'} & H & \xrightarrow{\psi'} & H'' & \to 0,
\end{array}$$

induces a fibre preserving map $\xi\colon (Y^K, p_1, X''^K) \to (Y_H{}^K, p_1, X''_H{}^K)$ by Lemma 6.4 when G', G, G'', H', and $H \in \mathfrak{D}$. Hence $\xi_\#$ commutes with $\delta_\#$ as well as

with $\phi_{\#}$ and $\psi_{\#}$. Thus a homomorphism of the exact coefficient sequence induces a homomorphism of the sequence corresponding to that exact coefficient sequence. This completes the proof.[4]

7. Proof of the universal coefficient theorem. This section is devoted to proving that the sequence

$$(*)\qquad 0 \to \pi^r(K,L) \otimes G \xrightarrow{\alpha} \pi^r(K,L;G) \xrightarrow{\beta} \operatorname{Tor}(\pi^{r+1}(K,L), G) \to 0$$

is exact under various hypotheses.

LEMMA 7.1. *$\alpha\colon \pi^r(K,L) \otimes G \to \pi^r(K,L;G)$ is natural with respect to maps $f\colon (K,L) \to (K',L')$, and if $G \varepsilon \mathfrak{D}$, then it is natural with respect to homomorphisms $\phi\colon G \to H$. Moreover, if G is finitely generated and free, then α is an isomorphism.*

Proof. The naturality statements are obvious. If $G = Z$, α is obviously an isomorphism because $\pi^r(K,L;Z) = \pi^r(K,L)$. Moreover, the functors $\pi^r(K,L) \otimes G$ and $\pi^r(K,L;G)$ are additive [5] on $\mathfrak{D}$ by the results of Section 3. Hence they commute with finite direct sums, and α is an isomorphism if G is finitely generated and free.

Proof of the first part of 3.6. Let $0 \to R \xrightarrow{i} F \xrightarrow{j} G \to 0$ be exact, where F is a finitely generated free abelian group. R is finitely generated and free also. By 3.5, the sequence corresponding to this coefficient sequence is exact (R and $F \varepsilon \mathfrak{D}$):

$$\cdots \to \pi^r(K,L;R) \xrightarrow{i_{\#}} \pi^r(K,L;F) \xrightarrow{j_{\#}} \pi^r(K,L;G) \xrightarrow{\delta_{\#}} \pi^{r+1}(K,L;R)$$
$$\xrightarrow{i_{\#}} \pi^{r-1}(K,L;F) \to \cdots.$$

Hence the following sequence is exact:

$$0 \to \operatorname{Ker}\delta_{\#} \to \pi^r(K,L;G) \to \operatorname{Im}\delta_{\#} \to 0.$$

However, $\operatorname{Ker}\delta_{\#} = \operatorname{Im} j_{\#} \approx \operatorname{Coker} i_{\#}$ and $\operatorname{Im}\delta_{\#} = \operatorname{Ker} i_{\#}$ by exactness. By 7.1, we see that

$$\operatorname{Coker} i_{\#} \approx \operatorname{Coker}(\pi^r(K,L) \otimes R \to \pi^r(K,L) \otimes F) \approx \pi^r(K,L) \otimes G$$

and

$$\operatorname{Ker} i_{\#} \approx \operatorname{Ker}(\pi^{r+1}(K,L) \otimes R \to \pi^{r+1}(K,L) \otimes F) = \operatorname{Tor}(\pi^{r+1}(K,L), G)$$

[4] The reader who is familiar with the techniques of Spanier and Whitehead can construct an alternate proof of Theorem 3.5 using those techniques.

(see [5] for the elementary properties of Tor needed in this proof). Under these isomorphisms, the inclusion $\operatorname{Ker}\delta_{\#} \to \pi^r(K;L;G)$ goes over to α: $\pi^r(K,L)\otimes G \to \pi^r(K,L;G)$, and $\delta_{\#}$ defines

$$\beta\colon \pi^r(K,L;G) \to \operatorname{Tor}(\pi^{r+1}(K,L),G).$$

Hence the sequence (*) is exact under the hypothesis of 3.6. (*) is natural with respect to maps $f\colon (K,L)\to (K',L')$ because of 7.1 and the fact that the exact sequence of 3.5 is natural. A homomorphism $\phi\colon G\to H$ gives rise to a commutative diagram

$$\begin{array}{ccccccccc}
0 & \to & R & \xrightarrow{i} & F & \xrightarrow{j} & G & \to & 0 \\
 & & \downarrow & & \downarrow & & \downarrow{\scriptstyle\phi} & & \\
0 & \to & R' & \xrightarrow{i'} & F' & \xrightarrow{j'} & H & \to & 0,
\end{array}$$

where R, R', F, and F' are finitely generated free abelian groups. Since R, R', F, F', and $G \varepsilon \mathfrak{D}$, by 3.5 we obtain a commutative diagram:

$$\begin{array}{ccccccccc}
\cdots \to & \pi^r(K,L;G) & \xrightarrow{\delta_{\#}} & \pi^{r+1}(K,L;R) & \xrightarrow{i_{\#}} & \pi^{r+1}(K,L;F) & \to \cdots \\
 & \downarrow{\scriptstyle\phi_{\#}} & & \downarrow & & \downarrow & \\
\cdots \to & \pi^r(K,L;H) & \xrightarrow{\delta_{\#}} & \pi^{r+1}(K,L;R') & \xrightarrow{i_{\#}} & \pi^{r+1}(K,L;F') & \to \cdots.
\end{array}$$

Hence the induced map $\operatorname{Tor}(\pi^{r+1}(K,L),G)\to \operatorname{Tor}(\pi^{r+1}(K,L),H)$ commutes with β. By 7.1, $\phi_{\#}$ commutes with α. This completes the proof of the first part of 3.6. Note the close similarity between this proof and the proof of Theorem 2.2A of [11].

An exact sequence $0\to A \xrightarrow{i} B \xrightarrow{j} C \to 0$ *splits* if there exists a homomorphism $k\colon C\to B$ such that $jk =$ the identity on C. For abelian groups, this is equivalent to the statement that $B = A + C$, the direct sum of A and C.

LEMMA 7.2. *The exact sequence (*) splits for $G = Z_p$, p an odd prime.*

Proof. By 5.3, $\pi^r(K,L;Z_p)$ is a vector space over Z_p and hence, $\pi^r(K,L;Z_p)$, as a group, is a direct sum of copies of Z_p. This implies that (*) splits when $G = Z_p$.

LEMMA 7.3. *If (K,L) is such that $\pi^r(K,L)$ is finitely generated, then the exact sequence (*) splits for $G = Z_{p^s}$, p an odd prime.*

Proof. Corresponding to the coefficient homomorphism $\phi\colon Z_{p^s} \to Z_p$ sending a generator into a generator, we have the commutative diagram (by 7.1):

$$\begin{array}{ccc} \pi^r(K,L)\otimes Z_{p^s} & \xrightarrow{1\otimes\phi} & \pi^r(K,L)\otimes Z_p \\ \downarrow{\scriptstyle \alpha_1} & & \downarrow{\scriptstyle \alpha} \\ \pi^r(K,L;Z_{p^s}) & \xrightarrow{\phi_\#} & \pi^r(K,L;Z_p). \end{array}$$

α and α_1 are monomorphisms. We can write

$$\pi^r(K,L) = Z + Z + Z_{p_1^{m_1}} + \cdots + Z_{p_k^{m_k}}$$

by hypothesis. We obtain generators $\{x_j\}$ of $\pi^r(K,L)\otimes Z_{p^s}$ from generators of Z and $Z_{p_i^{m_i}}$ for $p_i = p$. It is obvious that $(1\otimes\phi)(x_j) \neq 0$ for all j, and hence $\alpha(1\otimes\phi)(x_j) \neq 0$ for all j. However, if (*) does not split for $G = Z_{p^s}$, then some x_j is such that $\alpha_1(x_j)$ is divisible by p in $\pi^r(K,L;Z_{p^s})$. Hence $\phi_\#\alpha_1(x_j)$ is divisible by p in $\pi^r(K,L;Z_p)$, but each element of $\pi^r(K,L;Z_p)$ has order p. Thus $0 = \phi_\#\alpha_1(x_j) = \alpha(1\otimes\phi)(x_j) \neq 0$. This is a contradiction.

LEMMA 7.4. (*H-lemma*). *Given a commutative diagram where each row and each column is exact. If the first and third columns and the middle row split, then every row and every column splits.*

$$\begin{array}{ccccccccc} & & 0 & & 0 & & 0 & & \\ & & \downarrow & & \downarrow & & \downarrow & & \\ 0 & \to & A_1 & \xrightarrow{f_1} & B_1 & \xrightarrow{g_1} & C_1 & \to & 0 \\ & & \downarrow{\scriptstyle c_1} & & \downarrow{\scriptstyle d_1} & & \downarrow{\scriptstyle e_1} & & \\ 0 & \to & A_2 & \xrightarrow{f_2} & B_2 & \xrightarrow{g_2} & C_2 & \to & 0 \\ & & \downarrow{\scriptstyle c_2} & & \downarrow{\scriptstyle d_2} & & \downarrow{\scriptstyle e_2} & & \\ 0 & \to & A_3 & \xrightarrow{f_3} & B_3 & \xrightarrow{g_3} & C_3 & \to & 0 \\ & & \downarrow & & \downarrow & & \downarrow & & \\ & & 0 & & 0 & & 0 & & \end{array}$$

Proof. By hypothesis, there are homomorphisms $c_2^{-1}\colon A_3 \to A_2$, $e_2^{-1}\colon C_3 \to C_2$, and $g_2^{-1}\colon C_2 \to B_2$ such that $c_2c_2^{-1} = 1$, $e_2e_2^{-1} = 1$, and $g_2g_2^{-1} = 1$ on A_3, C_3, and C_2 respectively. Define $g_3^{-1}\colon C_3 \to B_3$ by $g_3^{-1} = d_2g_2^{-1}e_2^{-1}$, then $g_3g_3^{-1} = g_3d_2g_2^{-1}e_2^{-1} = e_2g_2g_2^{-1}e_2^{-1} = e_2e_2^{-1} = 1$, and hence the bottom row splits. Let f_3^{-1} be the other component of the direct sum decomposition, i.e. $f_3^{-1}\colon B_3 \to A_3$ is such that $1 = f_3f_3^{-1} + g_3^{-1}g_3$ on B_3. Define $d_2^{-1} = g_2^{-1}e_2^{-1}g_3 + f_2c_2^{-1}f_3^{-1}\colon B_3 \to B_2$. Then

$$\begin{aligned} d_2d_2^{-1} &= d_2g_2^{-1}e_2^{-1}g_3 + d_2f_2c_2^{-1}f_3^{-1} = g_3^{-1}g_3 + f_3c_2c_2^{-1}f_3^{-1} \\ &= g_3^{-1}g_3 + f_3f_3^{-1} = 1, \end{aligned}$$

and hence the middle column splits. Let f_2^{-1} and c_1^{-1} be the other components of the direct sum decompositions of the second row and the first column, i.e. $f_2^{-1}f_2 = 1$ and $c_1^{-1}c_1 = 1$ on A_2 and A_1 respectively. Define $f_1^{-1}\colon B_1 \to A_1$ by $f_1^{-1} = c_1^{-1}f_2^{-1}d_1$. Then $f_1^{-1}f_1 = c_1^{-1}f_2^{-1}d_1f_1 = c_1^{-1}f_2^{-1}f_2c_1 = c_1^{-1}c_1 = 1$, and and hence every row and every column splits.

LEMMA 7.5. *If the exact sequence* (*) *splits for* G' *and* G'', *then it splits for* $G' + G'' = G$, *where* G' *and* $G'' \varepsilon \mathfrak{D}$.

Proof. $0 \to G' \to G \to G'' \to 0$ is a split exact sequence. The following is a commutative diagram:

$$\begin{array}{ccccccc}
 & & 0 & & 0 & & 0 \\
 & & \downarrow & & \downarrow & & \downarrow \\
 & 0 \to & \pi^r(K,L)\otimes G' & \to & \pi^r(K,L)\otimes G & \to & \pi^r(K,L)\otimes G'' \to 0 \\
 & & \downarrow \alpha_1 & & \downarrow \alpha & & \downarrow \alpha_2 \\
 & 0 \to & \pi^r(K,L;G') & \to & \pi^r(K,L;G) & \to & \pi^r(K,L;G'') \to 0 \\
 & & \downarrow \beta_1 & & \downarrow \beta & & \downarrow \beta_2 \\
 & 0 \to & \mathrm{Tor}(\pi^{r+1}(K,L),G') & \to & \mathrm{Tor}(\pi^{r+1}(K,L),G) & \to & \mathrm{Tor}(\pi^{r+1}(K,L),G'') \to 0 \\
 & & \downarrow & & \downarrow & & \downarrow \\
 & & 0 & & 0 & & 0
\end{array}$$

The rows are exact and split because $\otimes$, Tor, and $\pi^r(K,L;G)$ are additive functors when $G\,\varepsilon\,\mathfrak{D}$. The columns are exact by above, and the first and thirds columns split by hypothesis. Hence by 7.4, the middle column splits.

The last part of 3.6 now follows immediately from 7.3 and 7.5.

In order to prove 3.7, we first establish a direct limit theorem. Let $G\,\varepsilon\,\mathfrak{D}$, and let $\{G^\alpha;\pi_\alpha{}^\beta\}$ be the direct system of finitely generated subgroups of G [6; Chapt. VIII]. Then $G \approx \mathrm{dir}_\alpha \lim G^\alpha$. Note that each $G^\alpha\,\varepsilon\,\mathfrak{D}$. $\{\pi^r(K,L;G^\alpha);(\pi_\alpha{}^\beta)_\#\}$ is obviously a direct system of groups for $r > (N+1)/2$. We define $\xi\colon \mathrm{dir}_\alpha \lim \pi^r(K,L;G^\alpha) \to \pi^r(K,L;G)$ by

$$\xi(\{[a]\}) = (\pi_\alpha)_\#([a]),$$

where $[a]\,\varepsilon\,\pi^r(K,L;G^\alpha)$, $\{[a]\}$ is the element in $\mathrm{dir}_\alpha \lim \pi^r(K,L;G^\alpha)$ represented by $[a]$, and $\pi_\alpha\colon G^\alpha \to G$ is the canonical homomorphism. Since $\pi_r = \pi_\beta\pi_\alpha{}^\beta$ and all $G^\alpha\,\varepsilon\,\mathfrak{D}$, $(\pi_\alpha)_\# = (\pi_\beta)_\#(\pi_\alpha{}^\beta)_\#$ and ξ is well defined (note that $G\,\varepsilon\,\mathfrak{D}$ is necessary here).

THEOREM 7.6. *If* (K,L) *is a finite CW-pair, then* ξ *is an isomorphism for* $r > (N+1)/2$. *This isomorphism is natural with respect to maps* $f\colon (K,L) \to (K',L')$ *and homomorphisms* $\phi\colon G \to H$.

Proof. Let X be an $X(G,r)$-space, and let $[a] \in \pi^r(K,L;G)$. Since K is compact, $a(K) \subset X$ is also compact. Thus $a(K)$ is contained in a finite subcomplex X' of X [7; p. 96]. Assume X is the standard $X(G,r)$-space constructed in the proof of 3.1. Then a finite subcomplex of X must be an $X(G^\beta,r)$-space for some β. Hence a defines a map $a'\colon (K,L) \to (X',x_0)$ such that $(\pi_\beta)_\#([a']) = [a]$. Hence ξ is an epimorphism. A similar argument shows ξ is a monomorphism because $K \times I$ is also a compact CW-complex. The naturality statements are obvious.

Theorem 3.7 now follows immediately from the exact sequence (*), 7.6, and the fact that direct limits preserve exactness, $\otimes$, and Tor [5].

The following example shows that the universal coefficient theorem of 3.6 and 3.7 does not hold in general, even if $G \in \mathfrak{D}$. Let $G = Q =$ the additive group of the rationals and let K be an $X(Q,n)$-space. Then by 3.3, $\pi^n(K;Q) \approx \mathrm{Hom}(Q,Q) = Q$ and $\pi^n(K;Z) \approx \mathrm{Hom}(Q,Z) = 0$. The sequence (*), if exact, would give here that

$$0 \to 0 \otimes Q \to Q \to \mathrm{Tor}(\pi^{n+1}(K),Q) \to 0$$

is exact. However $\mathrm{Tor}(\pi^{n+1}(K),Q) = 0$, and hence $Q = 0$. This is a contradiction.

8. Proof of Theorem 3.8. The proof of Theorem 3.8, as the theorem itself, is a generalization of the proof of Theorem 3.1 of [11]. The proof is based on the generalized cohomotopy exact couple; i.e. Z is replaced everywhere by G.

Proof of 3.8. As is the proof of 3.1 of [11], it suffices to prove that

$$H^{r+1}(K,L;G_{(1)}) \in \mathcal{C},\cdots,H^N(K,L;G_{(N-r)}) \in \mathcal{C}$$

for $r > \mathrm{Max}((N+1)/2, n - \alpha(\mathcal{C};G))$. Again $n - r < n - (n - \alpha(\mathcal{C};G)) = \alpha(\mathcal{C};G)$, hence $G_{(1)} \in \mathcal{C},\cdots,G_{(n-r)} \in \mathcal{C}$ by definition of $\alpha(\mathcal{C};G)$. Since $G_{(s)}$ is finitely generated by 3.6, we may use the universal coefficient theorem for cohomology (2.3A of [11]); i.e.

$$0 \to H^{r+s}(K,L) \otimes G_{(s)} \xrightarrow{\alpha} H^{r+s}(K,L;G_{(s)}) \xrightarrow{\beta} \mathrm{Tor}(H^{r+s+1}(K,L),G_{(s)}) \to 0$$

is exact. $G_{(s)} \in \mathcal{C}$ for $s \leqq n - r$, therefore $H^{r+s}(K,L;G_{(s)}) \in \mathcal{C}$ for $s \leqq n - r$. The proof now differs from that of 3.1 of [11]. We have yet to show that

$H^{r+s}(K, L; G_{(s)}) \varepsilon \mathscr{C}$ for $s > n-r$ from the assumption that $H^{r+s}(K, L; G) \varepsilon \mathscr{C}$ for $s > n-r$. Since

$$(G+G')_{(s)} \approx G_{(s)} + G'_{(s)},\ H^r(K, L; G+G') \approx H^r(K, L; G) + H^r(K, L; G'),$$

and G is finitely generated, it suffices to show this for $G = Z$ or $G = Z_{p^t}$. The case $G = Z$ is obvious as in the proof of 3.1 of [11]. We now consider the case $G = Z_{p^t}$. By 3.6, $(Z_{p^t})_{(s)}$ can have only a p-primary component. Hence it suffices to show that if $H^r(K, L; Z_{p^t}) \varepsilon \mathscr{C}$ for $r > n$, then $H^r(K, L; Z_{p^s}) \varepsilon \mathscr{C}$ for $r > n$ and all $s \geqq 1$ (we have changed the notation to simplify the rest of the proof).

If $t = 1$, then it is obvious by induction on s using the exact sequence

$$H^r(K, L; Z_{p^{s-1}}) \to H^r(K, L; Z_{p^s}) \to H^r(K, L; Z_p)$$

corresponding to the exact coefficient sequence

$$0 \to Z_{p^{s-1}} \to Z_{p^s} \to Z_p \to 0.$$

If $t > 1$, then let r_0 be the largest integer r such that $H^r(K, L; Z_{p^q}) \notin \mathscr{C}$ for some q with $1 \leqq q < t$, and assume $r_0 > n$. Then corresponding to the exact coefficient sequence

$$0 \to Z_{p^{t-q}} \to Z_{p^t} \to Z_{p^q} \to 0,$$

we have an exact sequence

$$H^{r_0}(K, L; Z_{p^t}) \to H^{r_0}(K, L; Z_{p^q}) \to H^{r_0+1}(K, L; Z_{p^{t-q}}).$$

However,

$$H^{r_0}(K, L; Z_{p^t}) \varepsilon \mathscr{C} \text{ and } H^{r_0+1}(K, L; Z_{p^{t-q}}) \varepsilon \mathscr{C},$$

hence $H^{r_0}(K, L; Z_{p^q}) \varepsilon \mathscr{C}$ which is a contradiction. Thus $r_0 \leqq n$ and $H^r(K, L; Z_p) \varepsilon \mathscr{C}$ for $r > n$ and as remarked above, this implies that $H^r(K, L; Z_{p^s}) \varepsilon \mathscr{C}$ for $r > n$ and all $s \geqq 1$. This completes the proof of 3.8.

9. Cohomotopy operations. We conclude our discussion of generalized cohomotopy groups by defining the concept of a universally defined cohomotopy operation analogous to universally defined homotopy operations [4]. We classify these operation and compute the classifying groups.

Let A and $B \varepsilon \mathfrak{D}$ throughout this section. A *cohomotopy operation* of type $(n, q; A, B)$ is a function $\theta\colon \pi^n(K, L; A) \to \pi^q(K, L; B)$, defined for every CW-pair (K, L) with $N \leqq \mathrm{Min}(2n-2, 2q-2)$, such that if $f\colon (K, L) \to (K', L')$, then $\theta f^{\#} = f^{\#}\theta\colon \pi^n(K', L'; A) \to \pi^q(K, L; B)$.

Let $n+1 \leqq 2q-2$, and let $[b] \varepsilon \pi^q(X, x_0; B)$, where X is an $X(A, n)$-space. If $[a] \varepsilon \pi^n(K, L; A)$ and $N \leqq \mathrm{Min}(2n-2, 2q-2)$, then define

$\theta_b([a]) = [ba] \varepsilon \pi^q(K, L; B)$. Clearly θ_b is a cohomotopy operation of type $(n, q; A, B)$. Thus we have a function χ: $\pi^q(X, x_0; B) \rightarrow$ the set of cohomotopy operations of type $(n, q; A, B)$, defined by $\chi([b]) = \theta_b$.

THEOREM 9.1. *If $n+1 \leqq 2q-2$, then χ is a 1-1 correspondence.*

Proof. Let $\iota \varepsilon \pi^n(X, x_0; A)$ denote the class of the identity map of (X, x_0) into (X, x_0). Let $a \varepsilon [a] \varepsilon \pi^n(K, L; A)$, $a: (K, L) \rightarrow (X, x_0)$. Then $a^{\#}(\iota) = [a]$. Hence if θ is a given cohomotopy operation of type $(n, q; A, B)$, then

$$\theta([a]) = \theta(a^{\#}(\iota)) = a^{\#}\theta(\iota) = \theta(\iota)[a] = \theta_b([a])$$

where $[b] = \theta(\iota) \varepsilon \pi^q(X, x_0; B)$. Thus χ is onto. If $\theta_b = \theta_{b'}$, then

$$[b] = [b]\iota = \theta_b(\iota) = \theta_{b'}(\iota) = [b']\iota = [b'].$$

Thus χ is 1-1.

COROLLARY 9.2. *If $n+1 \leqq 2q-2$, then each θ is a homomorphism.*

Proof. By 9.1, $\theta = \theta_b$ for some $[b] \varepsilon \pi^q(X, x_0; B)$. Let $[a]$ and $[a'] \varepsilon \pi^n(K, L; A)$, then

$$\theta_b([a] + [a']) = [b]([a] + [a']) = [ba] + [ba'] = \theta_b([a]) + \theta_b([a'])$$

by 2.3.

We now compute $\pi^q(X, x_0; B)$, where X is an $X(A, n)$-space.

THEOREM 9.3. (a) $\pi^n(X, x_0; B) \approx \operatorname{Hom}(A, B)$,

(b) $\pi^{n+1}(X, x_0; B) \approx \operatorname{Ext}(A, B)$,

(c) $\pi^q(X, x_0; B) = 0$ *for* $q > n+1$, *and*

(d) *if B is finitely generated and $q < n$, then*

$$0 \rightarrow \operatorname{Ext}(A, Z_{(n+1-q)} \otimes B + \operatorname{Tor}(Z_{(n-q)}, B)) \rightarrow \pi^q(X, x_0; B)$$
$$\rightarrow \operatorname{Hom}(A, Z_{(n-q)} \otimes B + \operatorname{Tor}(Z_{(n-q-1)}, B)) \rightarrow 0$$

is an exact sequence.

Proof. (a) follows from 3.3 and 3.4. (b) follows from 3.8 and the universal coefficient theorem for cohomology because X has dimension $n+1$. (c) is trivial because X has dimension $n+1$. If B is finitely generated, then $B_{(s)} \approx Z_{(s)} \otimes B + \operatorname{Tor}(Z_{(s-1)}, B)$ by Theorem 3.6.

$$
\begin{tikzcd}
 & \vdots \arrow[d] & & & \\
\cdots \arrow[r] & 0 \arrow[r] \arrow[d, "j'"] & \cdots & & \vdots \arrow[d] \\
\cdots \arrow[r] 0 \arrow[r, "\Delta'"] & \Gamma^{q,n} \arrow[r, "i'"] \arrow[d, "j'"] & H^{n+1}(X, x_0; B_{(n+1-q)}) \arrow[r, "\Delta'"] & & 0 \arrow[r] \arrow[d] & \cdots \\
\cdots \arrow[r] 0 \arrow[r, "\Delta'"] & \Gamma^{q,n-1} \arrow[r, "i'"] \arrow[d] & H^{n}(X, x_0; B_{(n-q)}) \arrow[r, "\Delta'"] & & 0 \arrow[r] \arrow[d] & \cdots \\
 & \vdots \arrow[d] & & & \vdots \\
\cdots \arrow[r] 0 \arrow[r, "\Delta'"] & \Gamma^{q,q} \arrow[r, "i'"] \arrow[d, "j'"] & 0 \arrow[r] & \cdots \\
\cdots \arrow[r] & \pi^q(X, x_0; B) \arrow[r] \arrow[d] & 0 \arrow[r] & \cdots \\
 & \vdots & & &
\end{tikzcd}
$$

Figure 3.

Consider figure 3, a portion of the generalized cohomotopy exact couple with $G = B$, $(K, L) = (X, x_0)$, and $N = n + 1$. It is clear from figure 3 that the sequence

$$0 \to \Gamma^{q,n} \to \Gamma^{q,n-1} \to H^n(X, x_0; B_{(n-q)})' \to 0$$

is exact. However,

$$\Gamma^{q,n} \approx H^{n+1}(X, x_0; B_{(n+1-q)}) \approx \operatorname{Ext}(A, B_{(n+1-q)}), \Gamma^{q,n-1} \approx \Gamma^{q,q} \approx \pi^q(X, x_0; B),$$

and $H^n(X, x_0; B_{(n-q)}) \approx \operatorname{Hom}(A, B_{(n-q)})$. Combined with the above, this completes the proof.

10. Generalized homotopy groups. We conclude this paper with some brief remarks on generalized homotopy groups.

As one might expect, there is a theory of homotopy groups with coefficients in G which is dual in an intuitive sense to the theory developed above. The results of Spanier and Whitehead [17] make this duality precise. By these results, we are led to consider spaces having only one non-vanishing cohomology group; in particular, we consider homotopy classes of maps of such a space into arbitrary spaces. The duality of [17] gives theorems dual to our thecrems on cohomotopy groups with coefficients in G. These theorems are only valid in the stable range. However, as in ordinary homotopy theory, there is a natural group structure defined outside of the stable range, and many of the theorems extend beyond the stable range. We leave the details of these results to the reader.

Appendix.

11. Proof of Lemma 3.4. *Proof of* 3.4.[5] For the properties of Ext needed in this proof, see [5]. There is a natural homomorphism

$$\chi\colon \operatorname{Hom}(A,B)\otimes Z_2 \to \operatorname{Hom}(A,B\otimes Z_2)$$

defined by $[\chi(f\otimes 1)](a) = f(a)\otimes 1$, where $f\in\operatorname{Hom}(A,B)$ and 1 is the non-zero element of Z_2. We first show that χ is an isomorphism if A is free. Let $A = \sum_i Z_i, i\in I$. Then

$$\operatorname{Hom}(A,B) = \operatorname{Hom}(\sum_i Z_i, B) \approx \prod_i \operatorname{Hom}(Z_i,B) = \prod_i B_i$$

and

$$\operatorname{Hom}(A,B\otimes Z_2) = \operatorname{Hom}(\sum_i Z_i, B\otimes Z_2) \approx \prod_i \operatorname{Hom}(Z_i, B\otimes Z_2) = \prod_i (B\otimes Z_2)_i.$$

Moreover, the natural homomorphism $(\prod_i B_i)\otimes Z_2 \to \prod_i (B\otimes Z_2)_i$ is an isomorphism (see exercise E-6 in [6; Chapt. V]). Hence χ is an isomorphism if A is free.

Let $0\to R\to F\to G\to 0$ be exact, where F and R are free abelian groups. By definition of Ext, $\operatorname{Hom}(F,H)\to\operatorname{Hom}(R,H)\to\operatorname{Ext}(G,H)\to 0$ is exact. Thus the following is a commutative diagram with the rows exact:

$$\begin{array}{ccccccccc}
\operatorname{Hom}(F,H)\otimes Z_2 & \to & \operatorname{Hom}(R,H)\otimes Z_2 & \to & \operatorname{Ext}(G,H)\otimes Z_2 & \to & 0 & \to & 0 \\
\downarrow & & \downarrow & & \downarrow & & \downarrow & & \downarrow \\
\operatorname{Hom}(F,H\otimes Z_2) & \to & \operatorname{Hom}(R,H\otimes Z_2) & \to & \operatorname{Ext}(G,H\otimes Z_2) & \to & 0 & \to & 0.
\end{array}$$

The first two and the last two vertical homomorphisms are isomorphisms, and hence by the five-lemma [6; p. 16], $\operatorname{Ext}(G,H)\otimes Z_2 \approx \operatorname{Ext}(G,H\otimes Z_2)$. Now by hypothesis, $0\to G\xrightarrow{\xi} G$ is exact, where $\xi(g) = 2g$. Thus

$$\operatorname{Ext}(G,H)\xrightarrow{\xi^*}\operatorname{Ext}(G,H)\to 0$$

is exact, where ξ^* is multiplication by 2. Hence every element of $\operatorname{Ext}(G,H)$ is divisible by 2 and $\operatorname{Ext}(G,H\otimes Z_2)\approx\operatorname{Ext}(G,H)\otimes Z_2 = 0$. This completes the proof.

PRINCETON UNIVERSITY.

---

[5] This proof was worked out with the help of D. A. Buchsbaum.

REFERENCES.

[1] M. G. Barratt, "Track groups I," *Proceedings of the London Mathematical Society*, vol. 5 (1955), pp. 71-106.

[2] ———, "Homotopy ringoids and homotopy groups," *Quarterly Journal of Mathematics*, vol. 5 (1954), pp. 271-290.

[3] A. L. Blakers and W. S. Massey, "The homotopy groups of a triad, I," *Annals of Mathematics*, vol. 53 (1951), pp. 161-205.

[4] ———, "Products in homotopy theory," *Annals of Mathematics*, vol. 58 (1953), pp. 295-324.

[5] H. Cartan and S. Eilenberg, *Homological algebra*, Princeton University Press, 1956.

[6] S. Eilenberg and N. E. Steenrod, *Foundations of algebraic topology*, Princeton University Press, 1953.

[7] P. J. Hilton, *An introduction to homotopy theory*, Cambridge University Press, 1953.

[8] W. S. Massey, "Exact couples in algebraic topology," I and II, III, IV, and V, *Annals of Mathematics*, vol. 56 (1952), pp. 363-396, vol. 57 (1953), pp. 248-286.

[9] J. C. Moore, "Some applications of homology theory to homotopy problems," *Annals of Mathematics*, vol. 58 (1953), pp. 325-350.

[10] ———, "On homotopy groups of spaces with a single non-vanishing homology group," *Annals of Mathematics*, vol. 59 (1954), pp. 549-557.

[11] F. P. Peterson, "Some results on cohomotopy groups," *American Journal of Mathematics*, vol. 78 (1956), pp. 243-258.

[12] J.-P. Serre, "Homologie singuliere des espaces fibrés," *Annals of Mathematics*, vol. 54 (1951), pp. 425-505.

[13] ———, "Groupes d'homotopie et classes de groupes abéliens," *Annals of Mathematics*, vol. 58 (1953), pp. 258-294.

[14] E. H. Spanier, "Borsuk's cohomotopy groups," *Annals of Mathematics*, vol. 50 (1949), pp. 203-245.

[15] ——— and J. H. C. Whitehead, "A first approximation to homotopy theory," *Proceedings of the National Academy of Sciences*, vol. 39 (1953), pp. 655-660.

[16] ——— and J. H. C. Whitehead, "The theory of carriers and S-theory," *Algebraic Geometry and Topology* (A Symposium in Honor of S. Lefschetz), Princeton University Press, 1956.

[17] ——— and J. H. C. Whitehead, "Duality in homotopy theory," *Mathematika*, vol. 2 (1955), pp. 56-80.

[18] J. H. C. Whitehead, "Combinatorial homotopy," I, *Bulletin of the American Mathematical Society*, vol. 55 (1949), pp. 213-245.

A NOTE ON THE INTERPOLATION OF SUBLINEAR OPERATIONS.*

By A. P. CALDERÓN and A. ZYGMUND.[1]

The purpose of this note is to give an extension of M. Riesz' interpolation theorem for linear operations to certain non-linear ones.

Let R be a measure space. This means that we have a set function $\mu(E)$, non-negative and countably additive, defined for some ('measurable') subsets E of R. For any measurable (with respect to μ) function f defined on R we write

$$\left(\int_R |f|^r d\mu\right)^{1/r} = \|f\|_{r,\mu} \qquad (0 < r < \infty),$$

and denote by $\|f\|_{\infty,\mu}$ the essential (with respect to μ) upper bound of $|f|$. The set of functions f such that $\|f\|_{r,\mu}$ is finite $(0 < r \leqq \infty)$ is denoted by $L^{r,\mu}$. If no confusion arises, we write $\|f\|_r$, L^r for $\|f\|_{r,\mu}$, $L^{r,\mu}$.

Let R_1 and R_2 be two measure spaces with measures μ and ν respectively. Let $h = Tf$ be a transformation of functions $f = f(u)$ defined (almost everywhere) on R_1 into functions $h = h(v)$ defined on R_2. The most important special case is when T is a *linear operation.* This means that if Tf_1 and Tf_2 are defined, and if α_1, α_2 are complex numbers, then $T(\alpha_1 f_1 + \alpha_2 f_2)$ is defined and

$$T(\alpha_1 f_1 + \alpha_2 f_2) = \alpha_1 Tf_1 + \alpha_2 Tf_2.$$

Let $r > 0$, $s > 0$. A linear operation $h = Tf$ will be said to be *of type* (r, s) if it is defined for each $f \varepsilon L^{r,\mu}$ and if

$$\|Tf\|_{s,\nu} \leqq M \|f\|_{r,\mu}, \tag{1}$$

where M is independent of f. The least value of M is called the (r, s) *norm* of the operation.

Denote by (α, β) points of the square

$$0 \leqq \alpha \leqq 1, \qquad 0 \leqq \beta \leqq 1. \tag{Q}$$

* Received September 9, 1955.

[1] The research resulting in this paper was supported in part by the office of Scientific Research of the Air Force under contract AF 18(600)-1111.

The Riesz interpolation theorem (in the form generalized by Thorin (see [1]-[6] of the References at the end of the note) asserts that if a linear operation $h = Tf$ is simultaneously of types $(1/\alpha_1, 1/\beta_1)$ and $(1/\alpha_2, 1/\beta_2)$, with norms M_1 and M_2 respectively, and if

$$(2) \qquad \alpha = (1-t)\alpha_1 + t\alpha_2, \quad \beta = (1-t)\beta_1 + t\beta_2, \qquad (0 < t < 1)$$

then T is also of type $(1/\alpha, 1/\beta)$, with norm

$$(3) \qquad M \leqq M_1^{1-t} M_2^{t}.$$

The significance of this theorem is by now widely recognized, and its applications are many. Riesz himself deduced the result, through appropriate passages to limits, from a theorem about bilinear forms, and in this argument the linearity of T plays an important role. The same can be said of other proofs. There are however a number of interesting operations which are not linear and to which therefore the theorem cannot be applied. For the sake of illustration we mention one of them, first considered by Littlewood and Paley (see [7]), which has important application in Fourier series.

Given any $f \varepsilon L(0, 2\pi)$, we consider the function $F(z)$ regular for $|z| < 1$, whose real part is the Poisson integral of f, and imaginary part is zero at the origin. The Littlewood-Paley function is

$$g(\theta) = \{ \int_0^1 (1-\rho) | F'(\rho e^{i\theta}) |^2 d\rho \}^{\frac{1}{2}}.$$

The operation $g = Tf$ is clearly not linear. It satisfies, however, the following relations

$$(4) \qquad | T(f_1 + f_2) | \leqq | Tf_1 | + | Tf_2 |,$$

$$(5) \qquad | T(kf) | = | k | \, | Tf |,$$

for any constant k.

There are other interesting non-linear operations which have the same properties and it may be of interest to study the problem of interpolation of such operations. This is the object of this note.

We begin with general definitions.

We call an operation $h = Tf$ *sublinear,* if the following conditions are satisfied:

(i) Tf is defined (uniquely) if $f = f_1 + f_2$, and Tf_1 and Tf_2 are defined;

(ii) For any constant k, $T(kf)$ is defined if Tf is defined;

(iii) Conditions (4) and (5) hold.

In view of (5) we may, as in the linear case, consider inequalities (1) and introduce the notions of the *type* and *norm* of a sublinear operation. In what follows, the functions f will be defined (almost everywhere) on a space R_1 with measure μ, and the $h = Tf$ on a space R_2 with measure ν.

THEOREM. *Let (α_1, β_1) and (α_2, β_2) be any two points of the square Q. Suppose that a sublinear operation $h = Tf$ is simultaneously of types $(1/\alpha_1, 1/\beta_1)$ and $(1/\alpha_2, 1/\beta_2)$ with norms M_1 and M_2 respectively. Let (α, β) be given by* (2). *Then T is also of type $(1/\alpha, 1/\beta)$, with norm M satisfying* (3).

Proof. We easily deduce from conditions (i), (ii), (iii) that, if Tf_1, $Tf_2, \cdots, Tf_n$ are defined so is $T\{n^{-1}(f_1 + \cdots + f_n)\}$ and

$$(6) \qquad | T\{(f_1 + f_2 + \cdots + f_n)/n\} | \leqq n^{-1}(| Tf_1 | + \cdots + | Tf_n |).$$

We may suppose that $\alpha_1 \leqq \alpha_2$. Thus

$$(7) \qquad \alpha_1 \leqq \alpha \leqq \alpha_2.$$

Consider any f in $L^{1/\alpha, \mu}$ and write $f = f_1 + f_2$, where f_1 equals f at the points at which $| f | \leqq 1$, and equals 0 elsewhere. By (7),

$$| f_1 |^{1/\alpha_1} \leqq | f_1 |^{1/\alpha} \leqq | f |^{1/\alpha},$$

so that $f_1 \varepsilon L^{1/\alpha_1}$ and Tf_1 is defined, by hypothesis. Similarly $f_2 \varepsilon L^{1/\alpha_2}$ and Tf_2 is defined. It follows from (i) that $Tf = T(f_1 + f_2)$ is defined. Our task is to show that the $(1/\alpha, 1/\beta)$ norm M of T is finite and satisfies (3).

We assume for the time being that $\alpha > 0$, $\beta < 1$. Take any $f \varepsilon L^{1/\alpha}$. Without loss of generality we may suppose that

$$\| f \|_{1/\alpha} = 1.$$

Clearly

$$(8) \qquad \| Tf \|_{1/\beta} = \sup_g \int_{R_2} | Tf | \cdot g d\nu,$$

where g is non-negative and satisfies $\| g \|_{1/(1-\beta)} = 1$. We may confine our attention to functions g which are simple (a function is called *simple* if it takes only a finite number of values and is distinct from 0 on a set of finite measure; simple functions are dense in every L^s, $0 < s < \infty$; and in L^∞, if the space has finite measure). We make one more assumption, of which we shall free ourselves later, namely that f is also simple. We fix f and g and consider the integral

$$(9) \qquad I = \int_{R_2} | Tf | \cdot g d\nu.$$

Let $c_1, c_2, \cdots, c_m$ be the distinct from 0 (and different from each other) values of f. Let E_k be the set in which $f = c_k$, and let $\chi_k = \chi_k(u)$ be the characteristic function of E_k. Similarly let $c'_1, c'_2, \cdots, c'_n$ be the different from 0 values of g, E'_l the set where $g = c'_l$, and $\chi_l = \chi_l(v)$ the characteristic function of E'_l. Hence $f = \sum |c_k| \epsilon_k \chi_k$, $g = \sum c'_l \chi'_l$, where $|\epsilon_k| = 1$ and $c'_l > 0$.

Let $\alpha(z)$ and $\beta(z)$ be the right sides of (2), with z for t. Consider the non-negative function

$$(10) \qquad \Phi(z) = \int_{R_2} |T(|f|^{\alpha(z)/\alpha} \operatorname{sign} f)| \, g^{(1-\beta(z))/(1-\beta)} \, d\nu \qquad (z = x + iy),$$

which reduces to I for $z = t$. We show that $\Phi(z)$ is continuous and $\log \Phi(z)$ is subharmonic, in the whole plane.

Since, for each z, $|f|^{\alpha(z)/\alpha} \operatorname{sign} f$ is simple, and so is in L^{1/α_1}, the function $T(|f|^{\alpha(t)/\alpha} \operatorname{sign} f)$ is in L^{1/β_1}, and in particular is integrable over the set where $g > 0$. Hence $\Phi(z)$ exists for each z.

We have

$$(11) \qquad \begin{aligned} \Phi(z) &= \int_{R_2} |T\{\sum |c_k|^{\alpha(z)/\alpha} \epsilon_k \chi_k\}| \, \{\sum c'_l{}^{(1-\beta(z))(1-\beta)} \chi_l\} d\nu \\ &= \sum \int_{E'_l} c'_l{}^{(1-\beta(z))/(1-\beta)} \, |T\{\sum |c_k|^{\alpha(z)/\alpha} \epsilon_k \chi_k\}| \, d\nu \\ &= \sum \int_{E'_l} T\{c'_l{}^{(1-\beta(z))/(1-\beta)} \sum |c_k|^{\alpha(z)/\alpha} \epsilon_k \chi_k\} \, | \, d\nu, \end{aligned}$$

and it is enough to show that each integral of the last sum is continuous and its logarithm is subharmonic. In proving this we shall make repeated use of the inequality $||Tf_1| - |Tf_2|| \leqq |T(f_1 - f_2)|$, which is a consequence of (4).

We therefore fix l and write

$$\psi_z = \sum_k c'_l{}^{(1-\beta(z))/(1-\beta)} |c_k|^{\alpha(z)/\alpha} \epsilon_k \chi_k, \qquad \Psi(z) = \int_{E'_l} |T\psi_z| \, d\nu.$$

Clearly

$$\begin{aligned} |\Psi(z + \Delta z) - \Psi(z)| &\leqq \int_{E'_l} |T(\psi_{z+\Delta z} - \psi_z)| \, d\nu \\ &\leqq \| T(\psi_{z+\Delta z} - \psi_z) \|_{1/\beta_1} \{\nu(E'_l)\}^{1-\beta_1} \\ &\leqq M_1 \| \psi_{z+\Delta z} - \psi_z \|_{1/\alpha_1} \{\nu(E'_l)\}^{1-\beta_1}, \end{aligned}$$

and since $\psi_{z+\Delta z} - \psi_z$ is zero outside $\cup E_k$ and tends to 0, uniformly in u, as $\Delta z \to 0$, the norm $\| \psi_{z+\Delta z} - \psi_z \|_{1/\alpha_1}$ tends to 0, and Ψ is continuous at z. Hence Φ is continuous.

It is very well known that $\log \Psi(z)$ is subharmonic if and only if $\Psi(z)e^{h(z)}$ is subharmonic for every harmonic $h(z)$. We fix a harmonic function $h(z)$, and denote by $H(z)$ the analytic function whose real part is $h(z)$. Since the problem is local, we may consider h and H in a given circle. Write

$$\psi^*_z = \psi_z e^{H(z)}, \qquad \Psi^*(z) = \Psi(z)e^{h(z)} = \int_{E'_l} | T\psi^*_z | \, d\nu.$$

We fix z, take a $\rho > 0$, and denote by $z_1, z_2, \cdots, z_p$ a system of points equally spaced over the circumference of the circle with center z and radius ρ. We have

$$\psi^*_z(u) = \lim_{p\to\infty} 1/p \sum_{j=1}^{p} \psi^*_{z_j}(u),$$

uniformly in u. Since

$$\int_{E'_l} | T(\psi^*_z - 1/p \sum_1^p \psi_{z_j}) | \, d\nu \leqq \| T(\psi^*_z - 1/p \sum_1^p \psi^*_{z_j} \|_{1/\beta_1} \{\nu(E'_l)\}^{1-\beta_1}$$
$$\leqq M_1 \| \psi^*_z - 1/p \sum_1^p \psi^*_{z_j} \|_{1/\alpha_1} \{\nu(E'_l)\}^{1-\beta_1}$$

the left side tends to 0 as $p \to \infty$. In particular, as $p \to \infty$ we have

$$\delta_p = \int_{E'_l} | \, | T\psi^*_z | - | T(1/p \sum_1^p \psi^*_{z_j}) | \, | \, d\nu \to 0,$$

$$\int_{E'_l} | T\psi^*_z | \, d\nu \leqq \delta_p + 1/p \sum_1^p \int_{E'_l} | T\psi^*_{z_j} | \, d\nu,$$

$$\Psi^*(z) \leqq \lim_{p\to\infty} 1/p \sum_1^p \Psi^*(z_j) = 1/(2\pi) \int_0^{2\pi} \Psi^*(z + \rho e^{it}) \, dt.$$

Hence $\Psi^*(z)$ is subharmonic.

We have therefore proved that $\Phi(z)$ is continuous in the whole plane and its logarithm is subharmonic. Moreover $\Phi(z)$ is bounded in every vertical strip of the plane, since from (11), (4) and (5) we deduce that

$$\Phi(z) \leqq \sum_{k,l} | c_k |^{\alpha(z)/\alpha} c'_l{}^{(1-\beta(z))/(1-\beta)} \int_{E'_l} | T(\chi_k) | \, d\nu.$$

Next we show that $\Phi \leqq M_1$ on the line $x = 0$, and $\Phi \leqq M_2$ on $x = 1$. It is enough to prove the first inequality. If $x = 0$, then

$$\Phi(z) \leqq \| T(|f|^{\alpha(z)/\alpha} \operatorname{sign} f) \|_{1/\beta_1} \| g^{(1-\beta_1)/(1-\beta)} \|_{1/(1-\beta_1)}$$
$$\leqq M_1 \| \, |f|^{\alpha(z)/\alpha} \operatorname{sign} f \|_{1/\alpha_1} \leqq M_1 \| \, |f|^{\alpha_1/\alpha} \|_{1/\alpha_1} = M_1.$$

Since $\log \Phi(z)$ is bounded above and subharmonic in the strip $0 \leqq x \leqq 1$,

and does not exceed $\log M_1$ and $\log M_2$ on the lines $x=0$ and $x=1$ respectively, an application of the Three-Line Theorem for subharmonic functions shows that

$$\log \Phi(z) \leqq (1-t) \log M_1 + t \log M_2$$

on the line $x=t$ and, in particular, $I=\Phi(t) \leqq M_1^{1-t} M_2^{t}$.

Summarizing results, we have proved that

$$\| Tf \|_{1/\beta} \leqq M_1^{1-t} M_2^{t} \| f \|_{1/\alpha} \tag{12}$$

for each simple f. We show that this holds for every $f \varepsilon L^{1/\alpha}$.

We fix such an f, and for each $m=1,2,\cdots$ consider the decomposition $f=f'_m+f''_m$, in which $f'_m=f$ wherever $|f| \leqq m$, and $f'_m=0$ elsewhere; hence $|f''_m|$ is either 0 or else greater than m. Let f_m be a simple function equal to 0 wherever $f'_m=0$ and such that $|f_m-f'_m|<1/m$ everywhere. Then

$$\big|\, |Tf| - |Tf_m| \,\big| \leqq |T(f-f_m)| \leqq |T(f'_m-f_m)| + |Tf''_m|. \tag{13}$$

If we show that each term on the right tends to 0 almost everywhere as m tends to $+\infty$ through a sequence of values, then the inequality (12) with f_m for f, will lead, by Fatou's lemma, to the inequality for f.

Now f'_m-f_m is in $L^{1/\alpha}$, and so also in L^{1/α_1}, since $|f'_m-f_m|<1/m$. It follows that

$$\| T(f'_m-f_m) \|_{1/\beta_1} \leqq M_1 \| f'_m-f_m \|_{1/\alpha_1} \leqq M_1 \| f'_m-f_m \|_{1/\alpha}^{\alpha_1/\alpha} \to 0. \tag{14}$$

as $m \to \infty$. Similarly

$$\| Tf''_m \|_{1/\beta_2} \leqq M_2 \| f''_m \|_{1/\alpha_2} \leqq M_2 \| f''_m \|_{1/\alpha}^{\alpha_2/\alpha} \to 0. \tag{15}$$

The inequalities (14) and (15) imply that there is a sequence of m tending to $+\infty$ and such that $|T(f'_m-f_m)|$ and $|Tf''_m|$ tend to 0 almost everywhere. This completes the proof of the theorem.

It remains however to consider the two extreme cases $\alpha=0$ and $\beta=1$, which we previously put aside. These two cases cannot occur simultaneously. If $\beta=1$, we replace the right side of (8) by $\int_{R_2} |Tf| \, d\nu$ and the function $\Phi(z)$ of (10) by

$$\int_{R_2} | T(|f|^{\alpha(z)/\alpha} \operatorname{sign} f) | \, d\nu.$$

After that the proof proceeds as before. If $\alpha=0$, then also $\alpha_1=\alpha_2=0$;

but it is immediately seen that whenever $\alpha_1 = \alpha_2$ the theorem is a corollary of Hölder's inequality.

MASSACHUSETTS INSTITUTE OF TECHNOLOGY AND
THE UNIVERSITY OF CHICAGO.

REFERENCES.

[1] M. Riesz, "Sur les maxima des formes billinéaires et sur les fonctionnelles linéaires," *Acta Mathematica*, vol. 49 (1926), pp. 465-497.

[2] G. O. Thorin, "An extension of a convexity theorem due to M. Riesz," *Kungl. Fysiografiska Saellskapets i Lund Forhaendliger*, 8 (1939), Nr. 14.

[3] ———, *Convexity theorems*, Uppsala, 1948, pp. 1-57.

[4] R. Salem, "Convexity theorems," *Bulletin of the American Mathematical Society*, vol. 55 (1949), pp. 851-860.

[5] J. D. Tamarkin and A. Zygmund, "Proof of a theorem of Thorin," *Bulletin of the American Mathematical Society*, vol. 50 (1944), pp. 279-282.

[6] A. P. Calderón and A. Zygmund, "On the theorem of Hausdorff-Young," *Annals of Mathematical Studies*, vol. 25 (1950), pp. 166-188.

[7] J. E. Littlewood and R. E. A. C. Paley, "Theorems on Fourier series and power series, I," *Journal of the London Mathematical Society*, vol. 6 (1931), pp. 230-233.

ON SINGULAR INTEGRALS.*

By A. P. Calderón and A. Zygmund.[1]

1. **Introduction.** In earlier work [1] we considered certain singular integrals arising in various problems of Analysis and studied some of their properties. Here we present a new approach to such integrals. Unlike the method used in [1] it is based on the theory of Hilbert transforms of functions of one variable, but otherwise it is simpler and yields most results obtained previously, under far less restrictive assumptions. Unfortunately some important cases ($f \varepsilon L$ for instance) seem to be beyond its scope. We have been unable to decide whether the corresponding theorems as presented in [1] can be likewise strengthened.

Our present results can be summarized in the theorems presented below.

Let $x, y, z, \cdots$ denote vectors in n-dimensional Euclidean space E_n, $|x|$ the length of x and $x' = x|x|^{-1}$. Consider the integral

1.1 $$\tilde{f}_\epsilon(x) = \int_{|x-y|>\epsilon} K(x,y) f(y)\, dy,$$

where dy denotes the element of volume in E_n.

Theorem 1. *If $K(x,y) = N(x-y)$, where $N(x)$ is a homogeneous function of degree $-n$, i.e. such that $N(\lambda x) = \lambda^{-n} N(x)$ for every x and $\lambda > 0$, and if $N(x)$ has in addition the following properties*

i) *$N(x)$ is integrable over the sphere $|x| = 1$ and its integral is zero,*

ii) *$N(x) + N(-x)$ belongs to $L \log^+ L$ on $|x| = 1$,*

then, if $f(x) \varepsilon L^p$, $1 < p < \infty$, $\tilde{f}_\epsilon(x)$ as defined in 1.1 converges to a limit $\tilde{f}(x)$ in the mean of order p, and pointwise almost everywhere as $\epsilon \to 0$. Furthermore $\bar{f}(x) = \sup_\epsilon |\tilde{f}_\epsilon(x)|$ belongs to L^p and $\|\bar{f}\|_p \leqq A \|f\|_p$, where A is a constant depending on p and K, and $\|f\|_p$ is the L^p norm of f.

The condition that $N(x) + N(-x)$ be in $L \log^+ L$ on $|x| = 1$ cannot be relaxed. For given a function $\phi(t)$ such that $\phi(t)/t \log t \to 0$ as $t \to \infty$

* Received September 9, 1955.

[1] The research resulting in this paper was supported in part by the office of Scientific Research of the Air Force under contract AF 18(600)-1111.

there exists a function satisfying i) such that $\phi[|N(x)+N(-x)|]$ is integrable on $|x|=1$ but whose Fourier transform is unbounded, so that even if the pointwise limit of $\tilde{f}_\epsilon(x)$ exists (as is the case of $f(x)$ continuously differentiable and vanishing outside a bounded set), no relationship of the form $\|\tilde{f}\|_2 \leqq A\|f\|_2$ holds.

The Fourier transform $M(x)$ of $N(x)$ is a homogeneous function of degree zero and can easily be shown to be given by the formula

$$M(x) = \int N(y')\,[i\tfrac{1}{2}\pi\operatorname{sg}\cos\theta + \log|\cos\theta|]\,d\sigma,$$

where θ is the angle between the unit vectors x' and y', and $d\sigma$ is the element of "area" of the sphere $|x|=1$ over which the integral is extended. It is the presence of the term $\log|\cos\theta|$ which makes the class $L\log^+ L$ the best possible. Since we merely want to indicate this fact we omit further details.

THEOREM 2. *If*

$$K(x,y) = N(x, x-y)$$

where $N(x,y)$ is homogeneous of degree $-n$ in y and

i) *for every x, $N(x,y)$ is integrable over the sphere $|y|=1$ and its integral is zero,*

ii) *for a $q>1$ and every x, $|N(x,y)|^q$ is integrable over the sphere $|y|=1$ and its integral is bounded,*

then the same conclusions as in Theorem 1 hold about $\tilde{\tilde{f}}_\epsilon(x)$, provided that $f\,\varepsilon\,L^p$ with $q/(q-1)\leqq p<\infty$.

The condition that $p \geqq q/(q-1)$ is essential. We shall show by means of an example that if $p<q/(q-1)$, then $\tilde{\tilde{f}}_\epsilon(x)$ need no longer be in L^p.

A third type of integrals suggested by the theory of spherical summability of Fourier integrals is the object of the next two theorems.

THEOREM 3. *If*

$$K(x,y) = N(x, x-y)\,\psi(|x-y|)$$

where $\psi(t)$ is a Fourier-Stieltjes transform, $N(x,y)$ is homogeneous of degree $-n$ in y and

i) $|N(x,y)| \leqq F(y)$ *where $F(y)$ is a homogeneous function of degree $-n$ integrable over $|y|=1$,*

ii) *$\psi(t)$ is an even function and $N(x,y)$ is odd in y, i.e. $N(x,y) = -N(x,-y)$, or $\psi(t)$ is odd and $N(x,y)$ is even in y,*

then the same conclusions as in Theorem 1 hold about $\tilde{f}_\epsilon(x)$.

THEOREM 4. *If $K(x,y)$ is the same as in the previous theorem with condition i) replaced by*

i′) *for some $q > 1$ and every x, $|N(x,y)|^q$ is integrable over the sphere $|y| = 1$ and its integral is bounded,*

then the same conclusions about $\tilde{f}_\epsilon(x)$ hold provided that $f \in L^p$, $q/(q-1) \leqq p < \infty$.

In the cases of Theorems 2, 3 and 4 we may also consider the transposed integral 1.1, that is $\int_{|x-y|>\epsilon} K(y,x) f(y)\, dy$. The convergence in the mean of this integral in an immediate consequence of those theorems. The pointwise convergence does not follow readily though, and at individual points the integral may actually diverge even if f is continuously differentiable and vanishes outside a bounded set.

A straightforward application of Theorem 3 will yield the following statement about spherical summability of Fourier integrals.

THEOREM 5. *If the number n of variables of f is odd and $f \in L^p$, $1 < p \leqq 2$, then the spherical means of order $\frac{1}{2}(n-1)$ of the Fourier integral representation of f converge to f in the mean of order p.*

Whether this theorem remains valid for n even is an open question.

Finally, we might also mention two generalizations of the maximal theorem of Hardy and Littlewood which are obtained using the same ideas. These extensions are needed in the proofs of Theorems 1 and 2.

THEOREM 6. *Let $K_\epsilon(x,y) = \epsilon^{-n} N(x-y) \psi(\epsilon^{-1}|x-y|)$ where $N(x)$ is a non-negative homogeneous function of degree zero, integrable over $|x| = 1$, and $\psi(t)$ is a non increasing function of the real variable t such that $\psi(|x|)$ is integrable in E_n. Then if $f \in L^p$, $1 < p \leqq \infty$, and*

$$f^*(x) = \sup_\epsilon \left| \int K_\epsilon(x,y) f(y)\, dy \right|,$$

f^* belongs to L^p and

$$\| f^* \|_p \leqq A \, \| f \|_p,$$

where A is a constant depending on N, p and ψ.

The case when $N(x) \equiv 1$ and $x(t)$ is the characteristic function of the interval $(0,1)$ is well known.

THEOREM 7. *If $K_\epsilon(x,y) = \epsilon^{-n}N(x, x-y)\psi(\epsilon^{-1}|x-y|)$ where $N(x,y)$ is homogeneous of degree zero in y, $|N(x,y)|^q$, $q>1$, is integrable over the sphere $|y|=1$ and its integral is bounded, and $\psi(t)$ is the same as in the previous theorem, then f^* as defined in Theorem 6 is in the same L^p class as f and $\|f^*\|_p \leqq A\|f\|_p$, provided that $q/(q-1) \leqq p < \infty$.*

2. We start by showing that the integral 1.1 is meaningful. In the cases of Theorems 1 and 3 this is not quite evident.

In either case we have $|K(x,y)| \leqq F(|x-y|)$, where $F(y)$ is a homogeneous function of degree $-n$, integrable over the sphere $|y|=1$, and thus it will be sufficient to show that

$$\int_{|x-y|>\epsilon} F(x-y)\,|f(y)|\,dy$$

is absolutely convergent for almost every x and any $\epsilon>0$.

Let y' be a unit vector, t a real number and S a full sphere in E_n of diameter d. Then the integral

2.1 $$\int_\Sigma F(y')\,dy' \int_S dx \int_\epsilon^\infty |t^{-1}f(x-ty')|\,dt,$$

where dy' is the element of area of the unit sphere Σ, is finite. In fact, the inner integral is less than or equal to

$$A\Big[\int_{-\infty}^{+\infty} |f(x-ty')|^p\,dt\Big]^{1/p},^{2}$$

where A depends on ϵ and p but not on f. Substituting this expression for the inner integral in 2.1 and applying Hölder's inequality to the integral over S we find that the latter is dominated by

$$A\,|S|^{(p-1)/p}\Big[\int_S dx \int_{-\infty}^{+\infty} |f(x-ty')|^p\,dt\Big]^{1/p}.$$

Now the integral with respect to x can be computed first along lines parallel to y' and then over the space of such lines, rendering evident that its value does not exceed $d\,\|f\|_p{}^p$. Thus the integral over S in 2.1 is a bounded function of y' and 2.1 is therefore finite. Hence

$$\int_\Sigma F(y')\,dy' \int_\epsilon^\infty |t^{-1}f(x-ty')|\,dt$$

[2] Throughout the rest of the paper the letter A will stand for a constant, not necessarily the same in each occurrence.

is finite for almost all x. But the last is nothing but the expression of $\int_{|x-y|>\epsilon} F(x-y)\,|f(y)|\,dy$ in polar coordinates with origin at x. In other words, for any $\epsilon > 0$, 1.1 is absolutely convergent for almost every x.

3. In this section we shall prove Theorems 3 and 4.

Let $g(t)$ be a function of the real variable t belonging to L^p in $-\infty < t < \infty$, $1 < p < \infty$, and let $\epsilon(s)$ be an arbitrary positive measurable function in $-\infty < s < \infty$. Then the integral

$$\int_{|s-t|>\epsilon(s)} g(t)/(s-t)\,dt$$

represents a function whose L^p norm does not exceed the L^p norm of g multiplied by a constant A which depends on p but not on g or the function $\epsilon(s)$ (see [1], Chapter II, Theorem 1). More generally, the same holds for

$$e^{irs}\int_{|s-t|>\epsilon(s)} e^{-irt}g(t)/(s-t)\,dt,$$

with the same constant as before. Thus if $\mu(r)$ is a function of bounded variation in $-\infty < r < \infty$ from Minkowski's integral inequality it follows that the L^p norm of the function of s given by

$$\int_{-\infty}^{+\infty} e^{irs}\Big[\int_{|s-t|>\epsilon(s)} e^{-irt}g(t)/(s-t)\,dt\Big]\,d\mu(r)$$

is not larger than the L^p norm of g multiplied by the constant A above and by the total variation of μ. Now interchanging the order of integration in the expression above (which we may) and observing that the function $\epsilon(s)$ is positive and measurable but otherwise arbitrary we conclude that if

3.1 $$\bar{g}(s) = \sup_{\epsilon} \Big| \int_{|s-t|>\epsilon} \{\psi(s-t)/(s-t)\}\,g(t)\,dt \Big|,$$

where $\psi(s) = \int_{-\infty}^{+\infty} e^{isr}\,d\mu(r)$, then $\|\bar{g}\|_p \leqq A\,V(\mu)\,\|g\|_p$, $V(\mu)$ being the total variation of μ and A being a constant which only depends on p.

Let now $f(x)$ be a given function of L^p, $1 < p < \infty$, in E_n, y' a unit vector, and define

3.2 $$\tilde{f}_\epsilon(x, y') = \int_{|t|>\epsilon} t^{-1} f(x - ty')\,\psi(t)\,dt,$$

3.3 $$\bar{f}(x, y') = \sup_{\epsilon} |\tilde{f}_\epsilon(x, y')|.$$

Clearly $\tilde{f}_\epsilon$ exists for almost all (x, y') and is a measurable function of (x, y'). Furthermore, for almost all (x, y') it is a continuous function of ϵ, so that if we restrict ϵ to rational values in 3.3 we obtain the same value for $\bar{f}$ almost everywhere in (x, y'), which shows that $\bar{f}$ is also measurable.

Now it is readily seen that $\bar{f}(x, y')$ restricted to any straight line parallel to y' is precisely the integral in 3.1 of the function $f(x)$ restricted to the same line. Consequently

$$\int_{-\infty}^{+\infty} \bar{f}(x - ty', y')^p \, dt \leqq A^p V(\mu)^p \int_{-\infty}^{+\infty} | f(x - ty') |^p \, dt,$$

and integrating this inequality over the space of lines parallel to y' we obtain

3.4 $$\int \bar{f}(x, y')^p \, dx \leqq A^p V(\mu)^p \int | f(x) |^p \, dx.$$

Define now

3.5 $$f^\dagger(x) = \tfrac{1}{2} \int_\Sigma \bar{f}(x, y') F(y') \, dy',$$

3.6 $$\tilde{f}_\epsilon(x) = \tfrac{1}{2} \int_\Sigma \tilde{f}_\epsilon(x, y') N(x, y') \, dy',$$

where F and N are the functions introduced in Theorem 3. On account of 3.3 it follows that $| \tilde{f}_\epsilon(x) | \leqq f^\dagger(x)$, and Minkowski's integral inequality applied to 3.5, and 3.4 gives

3.7 $$\| f^\dagger \|_p \leqq \tfrac{1}{2} A V(\mu) \int_\Sigma F(y') \, dy' \, \| f \|_p.$$

But the function $\tilde{f}_\epsilon(x)$ defined in 3.6 coincides with the integral 1.1 as specified in Theorem 3. To see this one merely has to substitute $\tilde{f}_\epsilon(x, y')$ for its value in 3.6 and observe that one obtains 1.1 in polar coordinates with origin at x. Interchanging the order of integration is permissible wherever 1.1 is absolutely convergent, that is, almost everywhere. Thus we have proved that under the assumptions of Theorem 3, $\bar{f}(x) = \sup_\epsilon | \tilde{f}_\epsilon(x) |$ belongs to L^p, and that $\| \bar{f} \|_p \leqq A \| f \|_p$. A more explicit estimate of the constant involved appears in the right-hand side of 3.7, where A depends only on p.

We now prove that the same holds under the assumptions of Theorem 4.

We redefine $f^\dagger(x)$ and $\tilde{f}_\epsilon(x)$ by means of the formulas

$$f^\dagger(x) = \tfrac{1}{2} \int_\Sigma \bar{f}(x, y') | N(x, y') | \, dy', \qquad \tilde{f}_\epsilon(x) = \tfrac{1}{2} \int_\Sigma \tilde{f}_\epsilon(x, y') N(x, y') \, dy'.$$

First we observe that the $\tilde{f}_\epsilon(x)$ just introduced coincides with the $\tilde{f}_\epsilon(x)$

in 1.1. For the last integral above is nothing but 1.1 expressed in polar coordinates with origin at x. On account of 3.3 it follows again that $|\tilde{f}_\epsilon(x)| \leqq f^\dagger(x)$, and Hölder's inequality and Fubini's theorem yield

$$\int f^\dagger(x)^p\,dx = 2^{-p}\int [\int_\Sigma \bar{f}(x,y')\,|N(x,y')|\,dy']^p\,dx$$

$$\leqq 2^{-p}\int [\int_\Sigma \bar{f}(x,y')^p\,dy'][\int_\Sigma |N(x,y')|^{p'}\,dy']^{p/p'}\,dx$$

$$\leqq 2^{-p}\int_\Sigma dy'[\int \bar{f}(x,y')^p[\int_\Sigma |N(x,y')|^{p'}\,dy']^{p/p'}\,dx],$$

where $p' = p/(p-1)$. Now $p' \leqq q$, so that condition i) in Theorem 4 implies that the innermost integral in the last expression above is bounded. Hence if $B^{p'}$ is an upper bound for this integral and ω is the "area" of the unit sphere in E_n, 3.4 yields $\| f^\dagger \|_p \leqq \frac{1}{2} A V(\mu)\omega^{1/p}B \| f \|_p$. Thus we find again that $\| \tilde{f}(x) \|_p \leqq A \| f \|_p$.

Now we can prove that $\tilde{f}_\epsilon(x)$ converges in the mean and pointwise almost everywhere. The argument clearly covers both Theorem 3 and Theorem 4.

Let $\rho(t)$ be an even and continuously differentiable function equal to 1 for $t=0$ and vanishing outside the interval $(-1,1)$. It is readily seen that the Fourier transform of $\psi(t)\rho(t)$ is bounded and integrable. Consider now the function equal to t^{-1} for $|t| > \epsilon$ and zero otherwise. An easy computation shows that its Fourier transform is bounded uniformly in ϵ and converges pointwise as $\epsilon \to 0$. Consequently it follows from Parseval's formula that

3.8 $$\int_{|t|>\epsilon} \psi(t)\rho(t)/t\,dt$$

converges as $\epsilon \to 0$. Thus under the hypotheses of either Theorem 3 or Theorem 4, the integral

$$\int_{|x-y|>\epsilon} K(x,y)\rho(|x-y|)\,dy$$

converges as $\epsilon \to 0$. To see this one merely has to compute this integral in polar coordinates and use the fact pointed out above that 3.8 converges.

Let now $f(x)$ be continuously differentiable and vanish outside a bounded set. Then

$$\tilde{f}_\epsilon(x) = \int_{|x-y|>\epsilon} K(x,y)f(y)\,dy = \int_{|x-y|>\epsilon} K(x,y)\,[f(y) - f(x)\rho(|x-y|)]\,dy$$

$$+ f(x)\int_{|x-y|>\epsilon} K(x,y)\rho(|x-y|)\,dy.$$

The integrand in the first integral on the right is absolutely integrable over E_n, and the second integral converges as $\epsilon \to 0$. Consequently $\tilde{f}_\epsilon(x)$ converges.

In the general case, given $f(x)$ in L^p and $\delta > 0$ there exists a continuously differentiable g vanishing outside a bounded set such that $f = g + h$ and $\| h \|_p < \delta$. Since

$$\tilde{f}_\epsilon(x) = \tilde{g}_\epsilon(x) + \tilde{h}_\epsilon(x), \quad | \tilde{h}_\epsilon(x) | \leqq \bar{h}(x)$$

and $\| \bar{h} \|_p \leqq A \| h \|_p < A\delta$, and since $\tilde{g}_\epsilon(x)$ converges everywhere,

$$\overline{\lim}\, \tilde{f}_\epsilon(x) - \underline{\lim}\, \tilde{f}_\epsilon(x) \leqq 2\bar{h}(x),$$

and this implies that $\tilde{f}_\epsilon(x)$ converges almost everywhere because $\bar{h}(x)$ has arbitrarily small L^p norm. Finally, since $\tilde{f}_\epsilon(x) \to \tilde{f}(x)$ almost everywhere and $| \tilde{f}_\epsilon(x) | \leqq \bar{f}(x)$, the theorem on dominated convergence yields $\| \tilde{f}_\epsilon - \tilde{f} \|_p \to 0$ as $\epsilon \to 0$.

Theorems 3 and 4 are thus established.

4. The proof of Theorems 6 and 7 is based on the same technique we used in the preceding section.

Let $f(t) \geqq 0$ be a function defined in $-\infty < t < \infty$, and $\phi(t) = \psi(t) t^{n-1}$, where $\psi(t)$ is the function introduced in Theorem 6. Then $\phi(t)$ is integrable in $(0, \infty)$. Set

$$f^*(s) = \sup_\epsilon \epsilon^{-1} \int_0^\infty f(s+t) \phi(t\epsilon^{-1})\, dt, \qquad \epsilon > 0,$$

$$F_s(t) = t^{-1} \int_0^t f(s+t)\, dt, \qquad t > 0,$$

and $G(s) = \sup_t F_s(t)$. Then integration by parts gives

$$\epsilon^{-1} \int_0^\infty f(s+t)\phi(t\epsilon^{-1}) dt = -\epsilon^{-1} \int_0^\infty t F_s(t) d\phi(t\epsilon^{-1}) \leqq -\epsilon^{-1} G(s) \int_0^\infty t d\phi(t\epsilon^{-1})$$

$$= G(s) \int_0^\infty \phi(t) dt,$$

and consequently $f^*(s) \leqq G(s) \int_0^\infty \phi(t)\, dt$.

Now, a theorem of Hardy and Littlewood (see [3], p. 244), asserts that if $f \in L^p$, $1 < p < \infty$, then $G \in L^p$ and $\| G \|_p \leqq A \| f \|_p$, where A depends on p only. Therefore $\| f^* \|_p \leqq A \| f \|_p$, A now depending on p and ϕ.

Let now $f(x) \geqq 0$ be a function from L^p, $1 < p < \infty$ in E_n and y' a unit vector. Define

$$f^*(x, y') = \sup_\epsilon \epsilon^{-1} \int_0^\infty f(x + ty') \phi(t\epsilon^{-1})\, dt.$$

Then, as we have shown above,

$$\int_{-\infty}^{+\infty} f^*(x+ty',y')^p\,dt \leqq A\int_{-\infty}^{+\infty} f(x+ty')^p\,dt,$$

where A depends only on p and ϕ. Integrating over the space of lines parallel to y' we obtain

4.1 $$\int f^*(x,y')^p\,dx \leqq A\int f(x)^p\,dx.$$

Under the assumptions of Theorem 6 we have

$$f^*(x)=\sup_\epsilon \left|\int K_\epsilon(x,y)f(y)dy\right| \leqq \sup_\epsilon \epsilon^{-n}\int N(x-y)\psi(|x-y|\epsilon^{-1})f(y)dy$$

$$=\sup_\epsilon \int_\Sigma N(y')[\epsilon^{-1}\int_0^\infty f(x+ty')\phi(t\epsilon^{-1})dt]dy' \leqq \int_\Sigma N(y')f^*(x,y')dy',$$

and an application of Minkowski's integral inequality and 4.1 yield

$$\| f^* \|_p \leqq A\int_\Sigma N(y')\,dy'\,\| f \|_p,$$

where A depends on p and ψ only.

On the other hand, under the assumptions of Theorem 7 we have

$$f^*(x)=\sup_\epsilon \left|\int K_\epsilon(x,y)f(y)dy\right|$$

$$\leqq \sup_\epsilon \epsilon^{-n}\int |N(x,x-y)|\,\psi(|x-y|\epsilon^{-1})f(y)dy$$

$$=\sup_\epsilon \int_\Sigma |N(x,y')|[\epsilon^{-1}\int_0^\infty f(x+ty')\phi(t\epsilon^{-1})dt]dy'$$

$$\leqq \int_\Sigma |N(x,y')|\,f^*(x,y')dy'.$$

Hence

$$\int f^*(x)^p\,dx \leqq \int [\int_\Sigma |N(x,y')|\,f^*(x,y')\,dy']^p\,dx$$

$$\leqq \int [\int_\Sigma f^*(x,y')^p\,dy'][\int_\Sigma |N(x,y')^{p'}dy']^{p/p'}\,dx,$$

where $p'=p/(p-1)$. But $p'\leqq q$ so that the integral of $|N(x,y')|^{p'}$ is bounded. If $B^{p'}$ is a bound for this integral and ω is the area of the unit sphere in E_n, interchanging the order of integration and applying 4.1 we obtain finally $\| f^* \|_p \leqq A\omega^{1/p}B\,\| f \|_p$, where A only depends on p and ψ.

The proof of Theorems 6 and 7 is thus completed.

Remark. The methods used so far still yield results under slightly less restrictive assumptions about the type of integrability of $f(x)$. For instance if f vanishes outside a bounded set and $|f|\log^+|f|$ is integrable, one can still prove that under the assumptions of either Theorem 3 or Theorem 4, $\tilde{f}_\epsilon(x)$ converges in the mean order 1 on any bounded set.

This result is needed in the next section but only in the special case when $K(x,y)$ is the kernel of M. Riesz (see below), and in this form it is also contained in Theorem 7, Chapter 1 of [1]. We may thus safely omit further details.

5. The proof of Theorems 1 and 2 in their full generality is more complicated. In special cases they are contained in Theorems 3 and 4 respectively. In fact, if in Theorem 1 the function $N(x)$ is such that $N(x)=-N(-x)$, then Theorem 1 reduces to Theorem 3 with $N(x,y)=N(y)$ and $\psi(t)=1$. Similarly, if $N(x,y)$ in Theorem 2 is such that $N(x,y)=-N(x,-y)$, then Theorem 2 reduces to Theorem 4 with the same N and $\psi(t)=1$.

Since it is always possible to decompose the functions $N(x)$ and $N(x,y)$ in the sum of two,

$$N(x)=N_1(x)+N_2(x),\qquad N(x,y)=N_1(x,y)+N_2(x,y),$$

where

$$N_1(x)=N_1(-x),\qquad N_2(x)=-N_2(-x),$$

and

$$N_1(x,y)=N_1(x,-y),\qquad N_2(x,y)=-N_2(x,-y),$$

we need only treat the cases $N(x)=N(-x)$, $N(x,y)=N(x,-y)$ and for this purpose we shall use the device of representing f as a singular integral with the kernel of M. Riesz. Our original integral 1.1 will then appear as an iterated integral to which we shall be able to apply the preceding results.

In what follows we shall use vector valued functions but we shall introduce no special notation for them. When talking about the L^p norm of a vector valued function we shall be meaning the L^p norm of its absolute value. The inner product of two vectors will be denoted by their symbols with a dot in-between.

The kernel R of M. Riesz is vector valued and odd

$$R(x)=\pi^{-\frac{1}{2}(n+1)}\Gamma(\tfrac{1}{2}n+\tfrac{1}{2})x\,|x|^{-n-1}.$$

If

$$g_\epsilon(x)=-\int_{|x-y|>\epsilon}R(x-y)f(y)\,dy,$$

and $f \varepsilon L^p$, $1 < p < \infty$, then $g_\epsilon(x)$ is a vector valued function which, as $\epsilon \to 0$, converges in the mean of order p to a function $g(x)$. This follows from Theorem 1 by applying it to each component. On the other hand, if

$$f_\epsilon(x) = \int_{|x-y|>\epsilon} R(x-y) \cdot g(y)\, dy,$$

where the integrand is the inner product of the vectors displayed, then $f_\epsilon(x)$ converges likewise to $f(x)$. That it converges to a function h is again a consequence of Theorem 1. That $h = f$ can be verified for f bounded and vanishing outside a bounded set by taking Fourier transforms (see [2]), whence the general case follows from the continuity in L^p of the linear operation taking f into h.

Let $\phi(t)$ be a continuously differentiable function of the real variable t, $t \geqq 0$, equal to zero in $(0, \frac{1}{4})$ and to 1 in $(\frac{3}{4}, \infty)$, and $F(x)$ a homogeneous function of degree $-n$, such that $F(x) = F(-x)$ and that $|F| \log^+ |F|$ is integrable on the sphere $|x| = 1$. Suppose in addition that the integral of $F(x)$ over $|x| = 1$ is zero and consider

$$\int_{|x-y|>\epsilon} R(x-y) F(y)\, dy, \tag{5.1}$$

$$\int_{|x-y|>\epsilon} R(x-y) F(y) \phi(|y|)\, dy. \tag{5.2}$$

Since $|R(x)| \leqq A |x|^{-n}$ the second integral is absolutely convergent. The first has a singularity at $y = 0$, but, owing to the fact that $R(x)$ is continuously differentiable if $x \neq 0$, it can be given a natural meaning if $|x| > \epsilon$ by integrating outside a small sphere with center at $y = 0$ and taking the limit of the value obtained as the radius of the small sphere tends to zero.

The properties of the integrals above which we need are summarized in the following

LEMMA. *Under the preceding assumptions, as $\epsilon \to 0$, 5.2 converges in the mean of order 1 on any compact set, and 5.1 converges on any compact set not containing the point $x = 0$. The corresponding limits, $F_2(x)$ and $F_1(x)$, are odd functions, i.e. $F_1(x) = -F_1(-x)$, $F_2(x) = -F_2(-x)$. The function $F_1(x)$ is homogeneous of degree $-n$, and, for $|x| \geqq 1$,*

$$|F_1(x) - F_2(x)| \leqq A \int_\Sigma |F(y')|\, dy' \, |x|^{-n-1}. \tag{5.3}$$

There exists a homogeneous function $G(x)$ of degree zero such that for $|x| \leqq 1$,

$$|F_2(x)| \leqq G(x), \tag{5.4}$$

and

5.5 $$\int_\Sigma G(y')\,dy' < \infty.$$

If for some q, $1 < q < \infty$, $\int_\Sigma |F(y')|^q\,dy' < \infty$, *then the inequalities*

5.6 $$\int_\Sigma |F_1(y')|^q\,dy' \leqq A \int_\Sigma |F(y')|^q\,dy',$$

5.7 $$\int_\Sigma G(y')^q\,dy' \leqq A \int_\Sigma |F(y')|^q\,dy'$$

hold, with the constants A depending on q but not on F, and the integral in 5.2 converges to its limit in the mean of order q.

That the functions $F_1(x)$ and $F_2(x)$, if existent, are odd is clear. That $F_1(x)$ is homogeneous of degree $-n$ is also clear. To see that 5.1 converges in the mean between two spheres of radii $\rho_1 < \rho_2$ we observe that the contributions to the integral from the sphere $|y| \leqq \frac{1}{2}\rho_1$ and from the exterior of the sphere $|y| = 2\rho_2$ is bounded, and to the integral extended over $\frac{1}{2}\rho_1 \leqq |y| \leqq 2\rho_2$ we may apply the remark of Section 4 and obtain immediately the desired result. On the other hand, an application of Theorem 4 to each component of the vector valued integral 5.1 gives that the integral of $|F_1(x)|^q$ extended to the region between two spheres with center at $x = 0$ is dominated by the q-th power of the right hand of 5.6. Since $F_1(x)$ is homogeneous, 5.6 follows. A similar argument applies to the integral 5.2, except that in this case it will not be necessary to exclude a neighborhood of $x = 0$.

For the difference between $F_1(x)$ and $F_2(x)$ we get the following estimates

$$\begin{aligned}|F_2(x) - F_1(x)| &\leqq \left| \int R(x-y)F(y)[\phi(|y|) - 1]\,dy \right| \\ &= \left| \int [R(x-y) - R(x)]F(y)[\phi(|y|) - 1]\,dy \right| \\ &\leqq \int_{|y| \leqq 3/4} |F(y)|\,|R(x-y) - R(x)|\,dy.\end{aligned}$$

Now, it is readily seen that for $|x| \geqq 1$ and $|y| \leqq \frac{3}{4}$ we have the inequality

$$|R(x-y) - R(x)| \leqq A\,|x|^{-n-1}\,|y|.$$

Substituting this in the preceding integral we obtain 5.3.

In order to prove 5.4, 5.5 and 5.7 we proceed as follows. First we observe that, owing to the fact that $F(y)\phi(|y|)$ vanishes in $|y| \leqq \frac{1}{4}$, $F_2(x)$

is continuous and bounded in $|x| \leqq \frac{1}{8}$, and that $A \int_\Sigma |F(y')|\, dy'$ with an appropriate constant A independent of F is an upper bound for $|F_2(x)|$ in this particular domain $|x| \leqq \frac{1}{8}$.

In $\frac{1}{8} \leqq |x| \leqq 1$ we have

$$\begin{aligned} |F_2(x)| &\leqq \phi(|x|)|F_1(x)| + |F_2(x) - \phi(|x|)F_1(x)| \\ &\leqq |F_1(x)| + \left| \int R(x-y)[\phi(|y|)F(y) - \phi(|x|)F(y)]dy \right. \\ &= |F_1(x)| + \left| \int [R(x-y) - \chi(|y|)R(x)]\,[\phi(|y|) - \phi(|x|)]F(y)dy \right|, \end{aligned}$$

where χ is the characteristic function of the interval $(0,1)$.

Now, one verifies easily that if $\frac{1}{8} \leqq |x| \leqq 1$, then

$$|R(x-y) - \chi(|y|)R(x)| \leqq A\, |y|^{\frac{1}{2}} \, |x-y|^{-n}.$$

On the other hand, since $\phi(t)$ has a bounded derivative,

$$|\phi(|x|) - \phi(|y|)| \leqq A\, |\, |x| - |y|\, | \leqq A\, |x-y|.$$

Thus in the last integral, substituting this we get

$$|F_2(x)| \leqq |F_1(x)| + A \int |y|^{\frac{1}{2}}\, |F(y)|\, |x-y|^{-(n-1)}dy,$$

and since $|x| \geqq \frac{1}{8}$, it follows that

$$|F_2(x)| \leqq A[\,|x|^n\, |F_1(x)| + |x|^{n-\frac{3}{2}} \int |y|^{\frac{1}{2}}\, |F(y)|\, |x-y|^{-(n-1)}dy].$$

The integral on the right represents a homogeneous function of degree $-n + \frac{3}{2}$, and $F_1(x)$ is homogeneous of degree $-n$. Hence the right-hand side of the inequality is a homogeneous function of degree zero. It remains only to show 5.5 and 5.7. The contribution of the term $|x|^n\, |F_1(x)|$ can be estimated either by the fact that $F_1(x)$ is integrable between two concentric spheres with center at the origin, or by 5.6. The contribution of the other term is estimated by splitting the integral in two, one extended over the set $\frac{1}{16} \leqq |y| \leqq 2$, the other over the complement of this set. The second integral is readily seen to be bounded by $A \int_\Sigma |F(y')|\, dy'$ or $A[\int_\Sigma |F(y')|^q\, dy']^{1/q}$ and an application of the theorem of Young (see [3], page 71, where the theorem is stated and proved in a special case; but the proof extends obviously to the most general situation) to the first shows that it represents a function of the same class as $F(x)$ in $\frac{1}{16} \leqq |x| \leqq 2$. But the function under consideration is homogeneous of degree zero and, consequently, an estimate for

its norm in $\frac{1}{2} \leqq |x| \leqq \frac{3}{2}$ gives an estimate for its norm on the sphere $|x| = 1$. Collecting results, 5.5 and 5.7 follow and the proof of our lemma is complete.

Let now $N(x)$ have the property that $N(x) = N(-x)$ and satisfy the conditions of Theorem 1. Then, by the lemma above,

$$N_1(x) = \lim_{\epsilon \to 0} \int_{|y-x|>\epsilon} R(x-y) N(y) \, dy$$

also satisfies the conditions of Theorem 1 and $N_1(x) = -N_1(-x)$. Furthermore, if

$$N_2(x) = \lim_{\epsilon \to 0} \int_{|x-y|>\epsilon} R(x-y) N(y) \phi(|y|) \, dy,$$

then, for $|x| \geqq 1$,

5.8 $$|N_2(x) - N_1(x)| \leqq A \, |x|^{-n-1},$$

and for $|x| \leqq 1$,

5.9 $$|N_2(x)| \leqq G(x),$$

where $G(x)$ is a homogeneous function of degree zero integrable over the sphere $|x| = 1$.

Consider now the vector valued function $g(x)$ and

$$f(x) = \lim_{\epsilon \to 0} \int_{|x-y|>\epsilon} R(x-y) \cdot g(y) \, dy.$$

As we already know, if $|g(x)| \,\varepsilon\, L^p$, $1 < p < \infty$, then $f(x) \,\varepsilon\, L^p$ and $\| f \|_p \leqq A \, \| g \|_p$. Furthermore the integral above converges in the mean of order p and every function $f(x)$ in L^p can be thus represented.

We shall prove the following identity:

5.10 $$\int N(x-y) \phi(|x-y| \epsilon^{-1}) f(y) \, dy = \epsilon^{-n} \int N_2(\epsilon^{-1}(x-y)) \cdot g(y) \, dy.$$

If g is continuously differentiable and vanishes outside a bounded set, then, on account of absolute integrability,

5.11 $$\begin{aligned} &\int N(x-y) \phi(|x-y| \epsilon^{-1}) \, dy \int_{|y-z|>\delta} R(y-z) \cdot g(z) \, dz \\ &= \int [\int_{|y-z|>\delta} N(x-y) \phi(|x-y| \epsilon^{-1}) R(y-z) \, dy] \cdot g(z) \, dz. \end{aligned}$$

Now, as $\delta \to 0$, by the lemma above and by changing variables, the inner integral on the right is seen to converge to $\epsilon^{-n} N_2((x-y)\epsilon^{-1})$ in the mean

of order 1 on any compact set. Therefore the right-hand side above converges to

$$\epsilon^{-n}\int N_2((x-y)\epsilon^{-1})\cdot g(y)\,dy.$$

On the other hand,

$$\int_{|y-z|>\delta} R(y-z)\cdot g(z)\,dz=\int_{|y-z|>\delta} R(y-z)\cdot[g(z)-g(y)]\,dz,$$

and on account of the continuous differentiability of g, the right-hand side is readily seen to converge uniformly as $\delta\to 0$ and to be independent of δ for $\delta<1$ and $|y|$ sufficiently large. Therefore, the left-hand side of 5.11 is seen to converge to

$$\int N(x-y)\phi(|x-y|\epsilon^{-1})f(y)\,dy.$$

Thus 5.10 is proved for g continuously differentiable and vanishing outside a bounded set. In the general case, given $g\,\varepsilon\,L^p$, we take a sequence of continuously differentiable functions g_k, each vanishing outside a compact set and such that $\|g_k-g\|_p\to 0$ and $\sum_1^\infty\|g_{k+1}-g_k\|_p<\infty$. If

$$f_k(x)=\lim_{\epsilon\to 0}\int_{|x-y|>\epsilon} R(x-y)\cdot g_k(y)\,dy,$$

then $\|f_k-f\|_p\to 0$ and $\sum_1^\infty\|f_{k+1}-f_k\|_p<\infty$. From the finiteness of the series $\sum\|g_{k+1}-g_k\|_p$ and $\sum\|f_{k+1}-f_k\|_p$ it follows that the functions $\bar{g}=|g_1|+\sum|g_{k+1}-g_k|$ and $\bar{f}=|f_1|+\sum|f_{k+1}-f_k|$ are finite almost everywhere and belong to L^p. Thus the sequences

$$g_k(x)=g_1(x)+\sum_1^{k-1}[g_{j+1}(x)-g_j(x)] \text{ and } f_k(x)$$

converge almost everywhere and are dominated in absolute value by $\bar{g}$ and $\bar{f}$ respectively. Now the considerations of Section 2 show that the integral

$$\int |N(x-y)|\,\phi(|x-y|\epsilon^{-1})\bar{f}(y)\,dy\leqq\int_{|x-y|>\epsilon/4}|N(x-y)|\,\bar{f}(y)\,dy$$

is finite for almost all x. On the other hand, on account of 5.8 we have

$$\int |N_2((x-y)\epsilon^{-1})|\,\bar{g}(y)\,dy\leqq\int_{|x-y|<\epsilon}|N_2((x-y)\epsilon^{-1})|\,\bar{g}(y)\,dy$$

$$+\int_{|x-y|>\epsilon}|N_1((x-y)\epsilon^{-1})|\,\bar{g}(y)\,dy+A\int_{|x-y|>\epsilon}|x-y|^{-n-1}\,\bar{g}(y)\,dy.$$

The first and last integrals on the right are absolutely convergent for almost all x owing to the absolute integrability of $N_2(x)$ in $|x| \leqq 1$, and the remaining integral, by the considerations of Section 2, is also finite for almost all x. Hence both

$$\int |N(x-y)| \phi(|x-y|\epsilon^{-1})\bar{f}(y)\,dy$$

and

$$\int |N_2((x-y)\epsilon^{-1})| \bar{g}(y)\,dy$$

are finite for almost all x. Consequently we can pass to the limit in

$$\int N(x-y)\phi(|x-y|\epsilon^{-1})f_k(y)\,dy = \epsilon^{-n}\int N_2((x-y)\epsilon^{-1}) \cdot g_k(y)\,dy,$$

and we find that in the general case 5.10 holds for almost all x.

Now the proof of Theorem 1 is nearly completed. We have

$$\begin{aligned}\tilde{f}_\epsilon(x) &= \int_{|x-y|>\epsilon} N(x-y)f(y)\,dy = \int N(x-y)\phi(|x-y|\epsilon^{-1})f(y)\,dy \\ &\qquad - \int_{|x-y|<\epsilon} N(x-y)\phi(|x-y|\epsilon^{-1})f(y)\,dy \\ &= \epsilon^{-n}\int N_2((x-y)\epsilon^{-1}) \cdot g(y)dy - \int_{|x-y|<\epsilon} N(x-y)\phi(|x-y|)\epsilon^{-1})f(y)dy,\end{aligned}$$

and on account of 5.8 and 5.9 we find that

$$\begin{aligned}|\tilde{f}_\epsilon(x)| &\leqq \left|\int_{|x-y|>\epsilon} N_1(x-y) \cdot g(y)dy\right| + \epsilon^{-n}\int_{|x-y|<\epsilon} G((x-y)/|x-y|)|\, g(y)|\, dy \\ &\qquad + A\epsilon^{-n}\int_{|x-y|>\epsilon} |(x-y)\epsilon^{-1}|^{-n-1}\, |g(y)|\, dy \\ &\qquad + A\epsilon^{-n}\int |N((x-y)/|x-y|)|\, |f(y)|\, dy,\end{aligned}$$

whence Theorem 1 applied to the first term on the right, and Theorem 6 applied to the remaining ones yield

$$\|\tilde{f}\|_p = \|\sup_\epsilon |\tilde{f}_\epsilon|\,\|_p \leqq A\,\|g\|_p + A\,\|f\|_p \leqq A\,\|f\|_p.$$

From this convergence in the mean and almost everywhere of $\tilde{f}_\epsilon(x)$ follows as in the proof of Theorems 3 and 4. Theorem 1 is thus proved.

The proof of Theorem 2 proceeds along similar lines but the differences justify its presentation. Let $K(x,y)$ be as specified in Theorem 2 with the additional property that $K(x,y) = K(x,-y)$ and define

$$K_1(x,y)=\lim_{\epsilon\to 0}\int_{|y-z|>\epsilon} K(x,z)R(y-z)\,dz,$$

$$K_2(x,y)=\lim_{\epsilon\to 0}\int_{|y-z|>\epsilon} K(x,z)\phi(|z|)R(y-z)\,dz.$$

Both K_1 and K_2 are odd functions in y, and K_1 satisfies the conditions of Theorem 2. Furthermore, for $|y|\geqq 1$,

5.12 $$|K_2(x,y)-K_1(x,y)|\leqq A\,|y|^{-n-1},$$

with A independent of x, and for $|y|\leqq 1$

5.13 $$|K_2(x,y)|\leqq G(x,y),$$

where G is homogeneous of degree zero in y and such that

5.14 $$\int_\Sigma G(x,y')^q\,dy'$$

is bounded. All this follows from our lemma. Let now $g(x)$ be a vector valued function in L^p, $p\geqq q/(q-1)$ and set

$$f(x)=\lim_{\epsilon\to 0}\int_{|x-y|>\epsilon} R(x-y)\cdot g(y)\,dy.$$

We shall prove the identity

5.15 $$\int K(x,x-y)\phi(|x-y|\,\epsilon^{-1})f(y)\,dy$$
$$=\epsilon^{-n}\int K_2(x,(x-y)\epsilon^{-1})\cdot g(y)\,dy.$$

If $g(y)$ is continuously differentiable and vanishes outside a bounded set this identity is proved the same way we proved 5.10. In the general case, given $g\in L^p$ we take a sequence of continuously differentiable functions g_n, each vanishing outside a bounded set, and tending to g in the mean of order p. Since both $K(x,x-y)\phi(|x-y|\,\epsilon^{-1})$ and $K_2(x,(x-y)\epsilon^{-1})$ as functions of y are of integrable power $p/(p-1)$, and since the functions f_n corresponding to the g_n converge to f in the mean of order p, a passage to the limit under the integral sign yields 5.15 in its full generality.

Now we have

$$\tilde f_\epsilon(x)=\int_{|x-y|>\epsilon} K(x,x-y)f(y)dy=\int K(x,x-y)\phi(|x-y|\,\epsilon^{-1})f(y)dy$$
$$-\int_{|x-y|<\epsilon} K(x,x-y)\phi(|x-y|\,\epsilon^{-1})f(y)dy$$
$$=\epsilon^{-n}\int K_2(x,(x-y)\epsilon^{-1})\cdot g(y)dy-\int_{|x-y|<\epsilon} K(x,x-y)\phi(|x-y|\,\epsilon^{-1})f(y)dy,$$

and on account of 5.12 and 5.13 we find that

$$| \tilde{f}_\epsilon(x)| \leqq | \int_{|x-y|>\epsilon} K_1(x, x-y) \cdot g(y)dy |$$
$$+ \epsilon^{-n} \int_{|x-y|<\epsilon} G(x, (x-y)/|x-y|)| g(y)| dy$$
$$+ A\epsilon^{-n} \int_{|x-y|>\epsilon} |(x-y)\epsilon^{-1}|^{-n-1} | g(y)| dy$$
$$+ A\epsilon^{-n} \int_{|x-y|<\epsilon} | K(x, (x-y)/|x-y|)| | f(y)| dy,$$

whence Theorem 2 applied to the first term on the right and Theorem 7 applied to the remaining ones yield

$$\| \tilde{f} \|_p = \| \sup_\epsilon | \tilde{f}_\epsilon(x) | \|_p \leqq A (\| g \|_p + \| f \|_p) \leqq A \| f \|_p.$$

The argument can now be completed as before.

We close this section by showing that Theorem 2 ceases to hold for functions in L^p, $p < q/(q-1)$.

Let $p < q/(q-1)$ and take an integer n and $\alpha > 0$ so that

$$1/\alpha \leqq p/(p-1) - q, \qquad n/\alpha = p/(p-1).$$

Define $f(x)$ in E_n as follows: $f(x) = 1$ for $|x| \leqq 1$ and $f(x) = 0$ otherwise. Set

$$K(x,y) = |x|^\alpha |y|^{-n} \text{ for } |x| \geqq 1 \text{ and } |x' + y'| \leqq 1/|x|,$$
$$K(x,y) = -|x|^\alpha |y|^{-n} \text{ for } |x| \geqq 1 \text{ and } |x' - y'| \leqq 1/|x|$$

and $K(x,y) = 0$ otherwise. Then one sees readily that

$$\int_\Sigma | K(x,y')|^q dy' \leqq A |x|^{\alpha q - n + 1} \qquad |x| \geqq 1$$

But

$$\alpha q - n + 1 = \alpha q - \alpha p/(p-1) + 1 = \alpha[q - p/(p-1)] + 1 \leqq 0,$$

and consequently the integral above is bounded. On the other hand,

$$| \tilde{f}(x)|^p \geqq A |x|^{(\alpha - n)p} \text{ as } |x| \to \infty,$$

and

$$(n-\alpha)p = \alpha p/(p-1) = n,$$

so that $\tilde{f} \notin L^p$.

Remark. In order to simplify our presentation as far as possible we have omitted to give explicit estimates for the constant A in the inequalities

$\| \bar{f} \|_p \leqq A \| f \|_p$. In the paper that follows this, though, we shall need to know more about A. If, in Theorem 1, $N(x)$ is such that

$$\int_\Sigma | N(y') |^q dy' < \infty, 1 < q < \infty,$$

then

$$A = A_{pq} [\int_\Sigma | N(y') |^q dy']^{1/q},$$

where A_{pq} depends on p and q but *not* on $N(x)$. The reader will have little difficulty in verifying this statement himself, by estimating step by step the constants in the preceding proofs.

6. In this last section we shall prove Theorem 5. We restrict ourselves to the case of three or more variables.

Let $f(x)$ be a function L^p, $1 < p \leqq 2$. Then $f(x)$ has a Fourier transform f given by

6.1 $$\hat{f}(x) = \lim_{r \to \infty} \int_{|y| < r} e^{i(x \cdot y)} f(y) \, dy,$$

the limit being understood as a limit in the mean of order $p/(p-1)$. The spherical means of order $\frac{1}{2}(n-1)$ of the Fourier integral of f are given by

6.2 $$\sigma_r(f, y) = (2\pi)^{-n} \int_{|x| < r} \hat{f}(x) e^{-i(x \cdot y)} (1 - | x |^2 r^{-2})^{\frac{1}{2}(n-1)} \, dx.$$

If we assume that f vanishes outside a compact set we may replace $\hat{f}$ by its value 6.1 and interchange the order of integration obtaining

6.3 $$\sigma_r(f, y) = (2\pi)^{-n} \int [\int_{|x| < r} e^{ix \cdot (z-y)} (1 - | x |^2 r^{-2})^{\frac{1}{2}(n-1)} \, dx] f(z) \, dz.$$

If we set

$$F(z) = \int_{|y| \leqq 1} e^{-i(y \cdot z)} (1 - | y |^2)^{\frac{1}{2}(n-1)} \, dy,$$

6.3 becomes

6.4 $$\sigma_r(f, y) = (2\pi)^{-n} r^n \int F[r(y-z)] f(z) \, dz.$$

Now the function $F(z)$ is in L^2 and bounded and consequently it belongs to L^q for every $q \geqq 2$. Therefore 6.4 can be extended to an arbitrary f in L^p by a passage to the limit.

Our next step will be to prove that

$$\| \sigma_r(f, y) \|_p \leqq A \| f \|_p.$$

For this purpose we shall show that

$$F(z) = \psi(|z|)\,|z|^{-n}$$

where $\psi(t)$ is an odd Fourier-Stieltjes integral. We take an orthogonal coordinate system in E_n whose first coordinate axis coincides with z and denote by t the corresponding coordinate. Then

$$|z|^n F(z) = |z|^n \int_{|y|\leqq 1} e^{-i(y\cdot z)}(1-|y|^2)^{\frac{1}{2}(n-1)}\,dy$$
$$= |z|^n\,\omega_{n-2}\int_{-1}^{+1} e^{i|z|t}\Big[\int_0^{(1-t^2)^{\frac{1}{2}}}[1-(t^2+s^2)]^{\frac{1}{2}(n-1)}\,s^{n-2}\,ds\Big]dt,$$

where ω_{n-2} denotes the "area" of the $n-1$ dimensional unit sphere.

Now by setting $s^2 = v(1-t^2)$ the inner integral on the right is readily seen to be equal to

$$\tfrac{1}{2}(1-t^2)^{n-1}\int_0^1 (1-v)^{\frac{1}{2}(n-1)}v^{\frac{1}{2}(n-3)}\,dv,$$

and thus

$$|z|^n F(z) = A\,|z|^n\int_{-1}^{+1} e^{i|z|t}(1-t^2)^{n-1}\,dt = \psi(|z|),$$

where

$$\psi(t) = At^n\int_{-1}^{+1} e^{ist}(1-s^2)^{n-1}\,ds.$$

Since n is odd, $\psi(t)$ is odd and an n-fold integration by parts shows that $\psi(t)$ is a Fourier-Stieltjes transform.

Thus 6.4 becomes

6.5 $$\sigma_r(f,y) = (2\pi)^{-n}\int \psi(r\,|y-z|)\,|y-z|^{-n}f(z)\,dz.$$

Since, for each r, $\psi(rt)$ is the Fourier-Stieltjes transform of a function whose total variation is independent of r, Theorem 3 applied to the preceding integral yields

6.6 $$\|\sigma_r(f,y)\|_p \leqq A\,\|f\|_p,$$

with A independent of r.

Suppose now that $f(x)$ has continuous derivatives of all orders and vanishes outside a bounded set. Then $\hat{f}$ is absolutely integrable and the inversion theorem for Fourier transforms shows that $\sigma_r(f,y)$ converges to $f(y)$ as $r\to\infty$. Furthermore $\sigma_r(f,y)$ is bounded uniformly in r and an inspection of 6.5 shows that

$$|\sigma_r(f,y)| \leqq A\,|y|^{-n}$$

for $|y|$ sufficiently large, regardless of the value of r. This makes it clear that

6.7 $$\| \sigma_r(f, y) - f(y) \|_p \to 0$$

as $r \to \infty$.

In the general case, given $f \varepsilon L^p$ and $\epsilon > 0$ there exists a function g with continuous derivatives of all orders and vanishing outside a bounded set such that $\| f - g \|_p < \epsilon$. Then

$$\| \sigma_r(f, y) - f(y) \|_p \leqq \| \sigma_r(g, y) - g(y) \|_p + \| \sigma_r(f - g, y) \|_p + \| f - g \|_p,$$

and according to 6.6 and 6.7 as $r \to \infty$ the right-hand side has an upper limit not exceeding $(A + 1)\epsilon$. Since ϵ is arbitrary, the proof is complete.

In closing this section we point out that if $\sigma_r(f, y)$ is defined directly by means of 6.4, then the theorem holds for $f \varepsilon L^p$ for any p, $1 < p < \infty$. We also remark that the same method of proof can be used to establish the corresponding result for other appropriate methods of spherical summation.

MASSACHUSETTS INSTITUTE OF TECHNOLOGY AND
THE UNIVERSITY OF CHICAGO.

REFERENCES.

[1] A. P. Calderón and A. Zygmund, "On the existence of certain singular integrals," *Acta Mathematica*, vol. 88 (1952), pp. 85-139.

[2] J. Horvàth, "Sur les fonctions conjuguées à plusieurs variables," Koninkligke Nederlandse Akademie van Wetenschappen. Indagationes Mathematicae ex Actis Quibus Titulis. *Proceedings of the Section of Sciences*, vol. 15, No. 1 (1953), pp. 17-29.

[3] A. Zygmund, *Trigonometrical series*, Warsaw (1935).

ALGEBRAS OF CERTAIN SINGULAR OPERATORS.*

By A. P. CALDERÓN and A. ZYGMUND.[1]

---

1. In this note we study composition of singular integral operators of a type we have considered in earlier work.

Let $x = (\xi_1, \xi_2, \cdots, \xi_n)$ denote either a point of Euclidean n-space of coordinates $\xi_1, \xi_2, \cdots, \xi_n$, or the vector from $0 = (0, 0, \cdots, 0)$ to $(\xi_1, \xi_2, \cdots, \xi_n)$, and $|x|$ its length, that is $(\xi_1^2 + \cdots + \xi_n^2)^{\frac{1}{2}}$.

If $K(x)$ is a homogeneous function of degree $-n$, i. e. such that

$$K(\lambda x) = \lambda^{-n} K(x)$$

for every x and every $\lambda > 0$, and if

$$(1.1) \qquad \int_\Sigma K(x)\, d\sigma = 0 \quad \text{and} \quad \int_\Sigma |K(x)|^p\, d\sigma < \infty$$

for some $p > 1$, the integral being taken over the unit sphere Σ, $|x| = 1$, and $d\sigma$ denoting the elements of "area" of Σ, then for $f \varepsilon L^r$

$$(1.2) \qquad \tilde{f}_\epsilon(x) = \int_{|x-y|>\epsilon} K(x-y) f(y)\, dy$$

converges pointwise almost everywhere and in the mean order r as ϵ tends to zero, and the operation of taking f into the limit $\tilde{f}$ of the integral above is continuous in L^r, and

$$(1.3) \qquad \| \tilde{f} \|_r \leqq A_{r,p} \left[\int_\Sigma |K(x)|^p\, d\sigma \right]^{1/p} \| f \|_r,$$

where A_{rp} is a constant depending on p and r only (see [3], remark to Section 5).

This result suggests studying composition of operators of the form

$$(1.4) \qquad \mathcal{K}(f) = \alpha f + \tilde{f},$$

where α is a complex constant.

We shall consider

i) the class $\mathfrak{A}$ of all operators with $K(x)$ in C^∞ in $|x| > 0$, that is with $K(x)$ possessing derivatives of all orders if $x \neq 0$;

---

* Received September 9, 1955.

[1] The research resulting in this paper was supported in part by the office of Scientific Research of the Air Force under contract AF 18(600)-1111.

ii) the class $\mathcal{A}_p$, $p > 1$, of all operators $\mathcal{K}$ for which

$$\| \mathcal{K} \|_p = | \alpha | + [\int_\Sigma | K(x) |^p d\sigma]^{1/p} < \infty. \tag{1.5}$$

Such classes are obviously closed under addition and multiplication by scalars. To show that they are also closed under composition (operator multiplication) is one of the purposes of this note. Moreover, we intend to prove that $\mathcal{A}_p$, when endowed with the norm (1.5), is a commutative semi-simple Banach algebra. This will follow from the inequality

$$\| \mathcal{K} \mathcal{H} \|_p \leqq A_p \| \mathcal{K} \|_p \| \mathcal{H} \|_p, \tag{1.6}$$

where A_p is a constant depending on p only.

Since all operators under consideration are continuous in L^2 and commute with translations, if $\mathfrak{F}(g)$ denotes the Fourier transform of g, and $f \varepsilon L^2$, we must have

$$\mathfrak{F}[\mathcal{K}(f)] = \mathfrak{F}(f) \mathfrak{F}(\mathcal{K}),^2 \tag{1.7}$$

where $\mathfrak{F}(\mathcal{K})$ is a bounded function which we shall call the *Fourier transform of* $\mathcal{K}$. Further, and according to (1.3) and (1.5), and assuming that the constant $A_{r,p}$ in (1.3) is $\geqq 1$,

$$| \mathfrak{F}(\mathcal{K}) | \leqq A_{2,p} \| \mathcal{K} \|_p. \tag{1.8}$$

As we shall see, $\mathfrak{F}(\mathcal{K})$ is actually a homogeneous function of degree zero, continuous in $x \neq 0$. If $\mathcal{K} \varepsilon \mathcal{A}$, then $\mathfrak{F}(\mathcal{K})$ is in addition of class C^∞ in $x \neq 0$. Conversely, every homogeneous function of degree zero possessing derivatives of all orders in $x \neq 0$ is the Fourier transform of an operator in $\mathcal{A}$.

Finally we shall prove that an operator in $\mathcal{A}$ or $\mathcal{A}_p$ has an inverse in the same class if and only if its Fourier transform does not vanish. Since we may identify homogeneous functions of degree zero with their restrictions to the sphere $| x | = 1$, we can translate the last statement into the language of Banach Algebras and assert that the space of maximal ideals of $\mathcal{A}_p$ is homeomorphic with the sphere $| x | = 1$.

For the convenience of the reader we summarize our results in the following formal statements.

---

[2] In our special case this known general statement also follows from the fact that the Fourier transform of $K_\lambda(x)$, defined to be $K(x)$ for $| x | > \lambda$ and zero otherwise, converges boundedly to a bounded function as $\lambda \to 0$ (see [2], pp. 89-91). For if $f \varepsilon L^2$, the Fourier transform of (1.1) converges in L^2 to the product of a bounded function depending on K only and the Fourier transform of f.

THEOREM 1. *If $\mathfrak{A}$ is the class of all operators defined in* i), *then $\mathfrak{A}$ is closed under addition and operator multiplication. The Fourier transform of an operator in this class is a homogeneous function of degree zero and of C^∞ in $x \neq 0$, and conversely every such homogeneous function is the Fourier transform of an operator in $\mathfrak{A}$. The Fourier transform of the product of two operators is the product of their Fourier transforms, and consequently an operator in $\mathfrak{A}$ has an inverse in $\mathfrak{A}$ if and only if its Fourier transform does not vanish. If $k(x)$ is the Fourier transform of the kernel $K(x)$, and $\beta(k)$ is the least common upper bound for the absolute value of k and of its derivatives up to order $n+1$ evaluated in $|x| \geqq 1$, then for $|x| = 1$ we have*

$$(1.9) \qquad |K(x)| \leqq A\beta(k),$$

where A is a constant independent of K.

THEOREM 2. *If $\mathfrak{A}_p$ is the class of all operators defined in* ii) *and endowed with the norm* (1.5), *then $\mathfrak{A}_p$ becomes a semisimple commutative Banach Algebra under operator multplication, and* (1.6) *holds for the norm of the product of two operators. Then Fourier product of an operator in $\mathfrak{A}_p$ is a homogeneous function of degree zero continuous in $|x| \neq 0$, and the Fourier transform of a product is the product of the Fourier transforms of the factors.*

The existence of inverses and the functional calculus of operators in $\mathfrak{A}_p$ is based on the following two theorems.

THEOREM 3. *Let $g(x)$ be a function defined on the sphere Σ ($|x| = 1$) which is locally a restriction to Σ of Fourier transforms of operators in $\mathfrak{A}_p$: that is, every $x_0 \varepsilon \Sigma$ is contained in a neighborhood where $g(x)$ coincides with the restriction to Σ of the Fourier transform of an operator in $\mathfrak{A}_p$. Then there exists a single operator $\mathfrak{H}$ in $\mathfrak{A}_p$ whose Fourier transform coincides with g at all points of Σ.*

THEOREM 4. *Let $g(x)$ be a function defined on the sphere Σ, which is locally an analytic function of the restriction $h(x)$ to Σ of the Fourier transform of an operator in $\mathfrak{A}_p$; that is, for every $x_0 \varepsilon \Sigma$ there exists a power series $\Sigma a_n Z^n$ with positive radius of convergence such that $g(x) = \Sigma a_n[h(x) - h(x_0)]^n$ for x in some neighborhood of x_0. Then $g(x)$ is a restriction to Σ of the Fourier transform of an operator in $\mathfrak{A}_p$.*

COROLLARY. *An operator in $\mathfrak{A}_p$ has an inverse in $\mathfrak{A}_p$ if and only if its Fourier transform does not vanish. The space of maximal ideals of $\mathfrak{A}_p$ is homeomorphic to the sphere Σ ($|x| = 1$).*

One thing here must be stressed.

If an operator in $\mathfrak{A}_p$ is thought of as acting in the space L^2 of square integrable functions, then the fact that its Fourier transform does not vanish implies immediately that there is a bounded operator in L^2 which is the inverse of the given one. The fact, however, that this operator is in $\mathfrak{A}_p$ is non-trivial, and this fact is, of course, the essence of the preceding corollary. A similar remark applies to Theorems 2, 3 and 4.

The content of Theorem 4 can be described briefly by saying that an analytic function of the Fourier transform $h(x)$ of an operator in $\mathfrak{A}_p$ is again the Fourier transform of an operator in $\mathfrak{A}_p$. This analytic function need not be single valued, and the values of $h(x)$ might even be allowed to go through branch points of the function, provided that the conditions of Theorem 4 are respected at such points x. For example, if the function $h(x)$ has a continuous square root and coincides locally with Fourier transforms of operators in $\mathfrak{A}_p$ at all points where $h(x)$ vanishes then $h(x)$ has a square root in $\mathfrak{A}_p$.

The preceding theorems apply immediately to systems of singular integral operators in $\mathfrak{A}$ or $\mathfrak{A}_p$. Such systems may be thought of as a convolution of a square singular matrix kernel with a vector function plus a numerical matrix applied to the same function. The condition of invertibility then becomes that the matrix of the corresponding Fourier transforms have a nonvanishing determinant.

2. With things organized as we have them here it will be convenient to study first operators in $\mathfrak{A}$. Once the basic facts about such operators are established and the validity of (1.6) is proved in this special case, everything else will be relatively simple.

The following partly standard notation will be sufficient for our purposes. We shall write $f \cdot g$ for the (absolutely convergent) integral of $f(x)\bar{g}(x)$, $f * g$ for the convolution of f and g, $\mathcal{F}(f)$ for the Fourier transform of f. We shall also write $g^\lambda(x) = \lambda^n g(\lambda x)$, and denote by $g_\lambda(x)$ the function equal to g if $|x| \geqq \lambda$ and to zero otherwise (the latter notation will apply to kernels only and will not conflict with the notation on the left side of (1.2)).

By Γ we shall denote the class of all functions g of C^∞ such that g and all its derivatives are $O(|x|^{-k})$ as $|x| \to \infty$, for each $k > 0$. The Fourier transform of a function in Γ is in Γ; this we easily see by differentiating under the integral sign and integrating by parts.

We shall call a function f *radial* if it only depends on $|x|$. Fourier transforms of radial functions are radial (see [1] page 67). By a *corradial* function on the other hand we shall mean a function which is orthogonal to

all radial functions, i.e. such that $f \cdot g = 0$ for all radial g. The fact just quoted clearly implies that Fourier transforms of corradial functions are corradial.

Homogeneous functions satisfying the first condition 1.1 are corradial in an obvious sense, and conversely every corradial homogeneous function satisfies that condition. We shall therefore refer to homogeneous functions satisfying 1.1 as *corradial homogeneous functions.*

The argument which follows is based on a certain representation of homogeneous functions of a given degree.

Suppose that $g(x)$ is a corradial function in Γ. Then $g(0) = 0$ and

$$\int_0^\infty g^\lambda(x)\lambda^{-n-1+r}d\lambda \tag{2.1}$$

converges absolutely for $r > -1$. Moreover it represents a corradial homogeneous function of degree $-r$. Differentiation under the integral sign shows that this function is of C^∞ in $x \neq 0$.

Conversely, every corradial homogeneous function $K(x)$ of degree $-r$ can thus be represented by setting $g(x) = K(x)\rho(|x|)$ where $\rho(t)$ has continuous derivatives of all orders, vanishes in a neighborhood of 0 and ∞ nad such that

$$\int_0^\infty \lambda^{-1}\rho(\lambda)\,d\lambda = 1. \tag{2.2}$$

Let $K(x)$ be corradial homogeneous of degree $-n$ and of C^∞ in $x \neq 0$. Then, if $x \neq 0$, $K_\lambda(x)$ converges to $K(x)$ as $\lambda \to 0$, and it is not difficult to prove (see [2], pp. 89-91) that also $\mathfrak{F}(K_\lambda)$ converges pointwise and boundedly to a limit which we shall denote by $\mathfrak{F}(K)$. Consequently, if $f \varepsilon L^2$, $K_\lambda * f$ converges in mean of order 2, and the Fourier transform of its limit is $\mathfrak{F}(f)\mathfrak{F}(K)$. Thus the Fourier transform $\mathfrak{F}(\mathcal{K})$ of the operator in (1.4) is precisely $\alpha + \mathfrak{F}(K)$.

We shall prove presently that this function is homogeneous of degree zero and of C^∞ in $x \neq 0$, and that conversely every function with such properties is of this form. This will imply immediately that $\mathfrak{A}$ is closed under operator multiplication, as we stated in Theorem 1.

Let $\rho(x)$ be a radial function of C^∞ such that $\rho(0) = 1$, $\rho(x) = 0$ for $|x| \geqq 1$, and let $f(x)$ be any function of C^∞ vanishing outside a bounded set. Then $K_\lambda \cdot \rho = 0$ and

$$\begin{aligned} \mathfrak{F}(f) \cdot \mathfrak{F}(K) &= \lim_{\lambda \to 0} \mathfrak{F}(f) \cdot \mathfrak{F}(K_\lambda) = \lim_{\lambda \to 0} (f \cdot K_\lambda) \\ &= \lim_{\lambda \to 0} [f - f(0)\rho] \cdot K_\lambda = [f - f(0)\rho] \cdot K, \end{aligned} \tag{2.3}$$

the last integral being absolutely convergent since $f(x)-f(0)\rho(x)$ vanishes at 0. We now represent $K(x)$ by the formula (2.1) with $r=n$ and obtain

$$[f-f(0)\rho]\cdot K=[f-f(0)\rho]\cdot\int_0^\infty\lambda^{-1}g^\lambda d\lambda=\int_0^\infty\lambda^{-1}[f-f(0)\rho]\cdot g^\lambda d\lambda,$$

the change of the order of integration being justified by absolute convergence.

Now since g^λ is corradial and ρ radial we have $g^\lambda\cdot\rho=0$, and since

$$\mathfrak{F}(g^\lambda)=\lambda^n\mathfrak{F}(g)^{\lambda^{-1}},$$

as seen by changing variables in the Fourier integral of g^λ, setting $\lambda^{-1}=\mu$ we may further write

$$\begin{aligned}\int_0^\infty\lambda^{-1}[f-f(0)\rho]\cdot g^\lambda d\lambda&=\int_0^\infty\lambda^{-1}(f\cdot g^\lambda)d\lambda=\int_0^\infty\mathfrak{F}(f)\cdot\mathfrak{F}(g^\lambda)\lambda^{-1}d\lambda\\&=\int_0^\infty\mathfrak{F}(f)\cdot\mathfrak{F}(g)^\mu\mu^{-n-1}d\mu=\mathfrak{F}(f)\cdot\int_0^\infty\mathfrak{F}(g)^\mu\mu^{-n-1}d\mu\end{aligned}\tag{2.4}$$

changes of the order of integration being again justified by the absolute convergence of the integrals involved.

From the equality of the left side of (2.3) and the right side of (2.4) we conclude that if

$$K(x)=\int_0^\infty\lambda^{-1}g^\lambda(x)\,d\lambda,\tag{2.5}$$

then

$$\mathfrak{F}(K)=\int_0^\infty\mathfrak{F}(g)^\lambda\lambda^{-n-1}d\lambda.\tag{2.6}$$

Since these integrals represent the most general corradial homogeneous functions of degrees $-n$ and 0 respectively, of C^∞ in $x\neq0$, we have proved *that* $\mathfrak{F}(\mathcal{K})$ is corradial homogeneous of degree zero and of C^∞ in $x\neq0$, and that conversely every function with these properties is an $\mathfrak{F}(\mathcal{K})$.

We now pass to the proof of (1.9). For this purpose we assume that $K(x)$ is represented as in (2.5), and that $|x|=1$. Then

$$\begin{aligned}|K(x)|&=\left|\int_0^\infty g(\lambda x)\lambda^{n-1}d\lambda\right|\\&\leqq\sup|g(y)|\int_0^1\lambda^{n-1}d\lambda+\sup|g(y)|\,|y|^{n+1}\int_1^\infty\lambda^{-2}d\lambda,\end{aligned}$$

and we only have to estimate $\sup|g(y)|$ and $\sup|g(y)|\,|y|^{n+1}$ in terms of $\mathfrak{F}(K)$.

Let $\hat{g}$ denote the Fourier transform of g, and let $\xi_1,\xi_2,\cdots,\xi_n$ be the coordinates of x, and $\eta_1,\cdots,\eta_n$ those of y. Then

$$g(x) = \int e^{2\pi i(x\cdot y)}\hat{g}(y)\,dy,$$

$$\xi_k^{n+1}g(x) = (2\pi i)^{-n-1}\int e^{2\pi i(x\cdot y)}(\partial^{n+1}/\partial\eta_k^{n+1})\hat{g}(y)\,dy.$$

By Hölder's inequality $|x|^{n+1} \leqq n^{\frac{1}{2}(n-1)}\sum|\xi_i|^{n+1}$, and thus

$$|g(x)|\,|x|^{n+1} \leqq (2\pi)^{-n-1}n^{\frac{1}{2}(n-1)}\sum_k \int |(\partial^{n+1}/\partial\eta_k^{n+1})\hat{g}(y)|\,dy,$$

$|g(x)| \leqq \int |\hat{g}(y)|\,dy$, and it only remains to estimate the integral on the right in terms of $\mathfrak{F}(K)$.

For this purpose we set, as we may, $\hat{g}(y) = \mathfrak{F}(K)\rho(|y|)$, where $\rho(\lambda)$ is of C^∞, vanishes outside $1 \leqq \lambda \leqq 2$ and satisfies (2.2). This function we choose once for all independently of the particular kernel K under consideration. This makes it clear that $\hat{g}(y)$ and its derivatives can be estimated in terms of the derivatives of $\mathfrak{F}(K)$ and $\mathfrak{F}(K)$ itself. Furthermore $\hat{g}(y)$ vanishes outside $1 \leqq |y| \leqq 2$, and this fact makes it possible to estimate the integral above in terms of $\beta[\mathfrak{F}(K)]$. Collecting estimates we obtain (1.9).

Finally, from (1.9) we readily obtain

$$\|\mathcal{K}\|_p \leqq A\beta[\mathfrak{F}(\mathcal{K})] \tag{2.8}$$

for any operator $\mathcal{K}$ in $\mathfrak{A}$, A being a constant independent of p and $\mathcal{K}$, but not necessarily the same as in (1.9).

3. In this section we prove (1.6) for operators in $\mathfrak{A}$.

Let K and H be two corradial homogeneous functions of degree $-n$ of C^∞ in $x \neq 0$, and consider

$$K * H_\lambda = \lim_{\mu\to 0} K_\mu * H_\lambda. \tag{3.1}$$

As $\lambda \to 0$, this function converges pointwise for $x \neq 0$, and its limit J is a homogeneous function of degree $-n$. On the other hand, the Fourier transform of (3.1) converges to $\mathfrak{F}(K)\mathfrak{F}(H)$. We will show that J is corradial and that

$$\mathfrak{F}(K)\mathfrak{F}(H) = \alpha + \mathfrak{F}(J), \tag{3.2}$$

where α is a constant, and that

$$\|\mathcal{J}\|_p \leqq A_p\|\mathcal{K}\|_p\|\mathcal{H}\|_p, \tag{3.3}$$

$$|\alpha| \leqq A_p\|\mathcal{K}\|_p\|\mathcal{H}\|_p, \tag{3.4}$$

$\mathcal{H}$, $\mathcal{J}$, $\mathcal{K}$ being the convolution operators with kernels H, J, K, respectively.

First we easily see that $H_\lambda = (H_1)^{\lambda^{-1}}$; $K * H_\lambda = (K * H_1)^{\lambda^{-1}}$. Thus $K * H_\lambda - J_\lambda = (K * H_1 - J_1)^{\lambda^{-1}}$. Since H_1 is in L^2 the same holds for $K * H_1$. It follows that $K * H_1 - J_1$ is integrable over bounded sets. On the other hand it is not difficult to see that $K * H_1 - J_1$ is of order $|x|^{-n-1}$ as $|x| \to \infty$. Hence $K * H_1 - J_1$ is absolutely integrable.

Let now f be a function in Γ. A change of variables gives

$$(K * H_\lambda - J_\lambda) \cdot f = (K * H_1 - J_1)^{\lambda^{-1}} \cdot f = (K * H_1 - J_1) \cdot \lambda^{-n} f^\lambda$$

As $\lambda \to 0$, $\lambda^{-n} f^\lambda$ tends to $f(0)$ while remaining bounded, and this implies that $(K * H_\lambda - J_\lambda) \cdot f$ converges as $\lambda \to 0$. Now

$$K * H_\lambda \cdot f = \mathfrak{F}(K)\, \mathfrak{F}(H_\lambda) \cdot \mathfrak{F}(f)$$

also converges as $\lambda \to 0$, and consequently the same holds for $J_\lambda \cdot f$. But if $f(0) \neq 0$, $J_\lambda \cdot f$ cannot converge unless J is corradial. Hence J is corradial and $\mathfrak{F}(J_\lambda)$ converges boundedly to a limit $\mathfrak{F}(J)$.

Suppose now that $f(0) = 0$. Then

$$\begin{aligned}[\mathfrak{F}(K)\mathfrak{F}(H) - \mathfrak{F}(J)] \cdot \mathfrak{F}(f) &= \lim_{\lambda \to 0} [\mathfrak{F}(K)\mathfrak{F}(H_\lambda) - \mathfrak{F}(J_\lambda)] \cdot \mathfrak{F}(f) \\ &= \lim_{\lambda \to 0} (K * H_\lambda - J_\lambda) \cdot f.\end{aligned}$$

Since $f(0) = 0$, the last limit is zero, as we pointed out above. Consequently, if $g = \mathfrak{F}(f)$ we have

$$[\mathfrak{F}(K)\mathfrak{F}(H) - \mathfrak{F}(J)] \cdot g = 0$$

for any $g \in \Gamma$ with vanishing integral, and this is possible only if $\mathfrak{F}(K)\mathfrak{F}(H) - \mathfrak{F}(J)$ is a constant.

Next we estimate $\| \mathcal{J} \|_p$. First we note that, on account of homogeneity,

$$[\int |J_\lambda(x)|^p \, dx]^{1/p} = [n(p-1)\lambda^{(p-1)n}]^{-1/p} \| \mathcal{J} \|_p, \tag{3.5}$$

and similarly for H and K. Next, for $|x| \geqq 2$ we have

$$\begin{aligned}J_2(x) &= [(K - K_1) + K_1] * [(H - H_1) + H_1] \\ &= (K - K_1) * (H - H_1) + (K - K_1) * H_1 + (H - H_1) * K_1 + H_1 * K_1,\end{aligned}$$

and since the first term in the last sum vanishes for $|x| \geqq 2$, we see that

$$J_2(x) = K * H_1 + H * K_1 - H_1 * K_1$$

for $|x| \geqq 2$. Now (3.5) and (1.3) applied to this inequality yield (3.3). And from (1.8), (3.3) and (3.2) we easily derive (3.4).

It is clear that (3.3) and (3.4) imply (1.6).

4. The extension of (1.6) to operators in $\mathfrak{A}_p$ is straightforward.

Given two operators $\mathcal{H}$ and $\mathcal{K}$ in $\mathfrak{A}_p$, we take two sequences of operators $\mathcal{H}_n$, $\mathcal{K}_n$ in $\mathfrak{A}$ such that

$$\| \mathcal{H}_n - \mathcal{H} \|_p \to 0; \qquad \| \mathcal{K}_n - \mathcal{K} \|_p \to 0.$$

Then from the validity of (1.6) for operators in $\mathfrak{A}$ it follows that $\mathcal{H}_n\mathcal{K}_n$ is a Cauchy sequence in $\mathfrak{A}_p$, and therefore converges to a limit $\mathcal{J}$ in $\mathfrak{A}_p$ for which the inequality $\| \mathcal{J} \|_p \leqq A_p \| \mathcal{H} \|_p \| \mathcal{K} \|_p$ holds. Consequently, if we show that $\mathcal{J} = \mathcal{H}\mathcal{K}$ we will have shown that $\mathfrak{A}_p$ is closed under multiplication (composition) and that (1.6) holds for the product.

Consider (1.3) and (1.5). Assuming, as we may, that $A_{r,p} \geqq 1$, we see that $\mathcal{K}(f) = \alpha f + \tilde{f}$ satisfies $\| \mathcal{K}(f) \|_r \leqq A_{r,p} \| \mathcal{K} \|_p \| f \|_r$. Consequently the operator norm of $\mathcal{K}$ as an operator in L^r, which is defined as

$$\sup_f \| \mathcal{K}(f) \|_r / \| f \|_r,$$

is dominated by $A_{r,p} \| \mathcal{K} \|_p$. Since $\mathcal{H}_n \to \mathcal{H}$ and $\mathcal{K}_n \to \mathcal{K}$ in $\mathfrak{A}_p$, the same holds in the operator topology, and consequently $\mathcal{H}_n\mathcal{K}_n \to \mathcal{H}\mathcal{K}$ in the operator topology. On the other hand, $\mathcal{H}_n\mathcal{K}_n \to \mathcal{J}$ in $\mathfrak{A}_p$, and consequently the same holds in the operator topology. Hence $\mathcal{J} = \mathcal{H}\mathcal{K}$ and the proof is completed.

This also completes the proof of Theorem 2 since the fact that the Fourier transforms of operators in $\mathfrak{A}_p$ are continuous homogeneous functions of degree zero follows readily from (1.8) and the fact that $\mathfrak{A}$ is dense in $\mathfrak{A}_p$ and its elements have continuous Fourier transforms.

5. We now proceed to prove Theorems 3 and 4.

We might observe here that if we knew already that every maximal ideal in $\mathfrak{A}_p$ is the set of all operators whose Fourier transforms vanish at a point of the unit sphere Σ ($|x| = 1$), then Theorems 3 and 4 would merely be standard facts from Banach Algebras. In our present setup though we can prove Theorems 3 and 4 directly with comparatively little additional effort and obtain the structure of the maximal ideals as a consequence.

For simplicity of notation we shall denote the Fourier transform of an operator $\mathcal{H}$ in $\mathfrak{A}_p$ by h. The symbol $\| h \|_p$ will now stand for $\| \mathcal{H} \|_p$, and $\beta(h)$ will denote, as in Theorem 1 or Section 2, the least upper bound for the absolute value of h and its derivatives of order $n+1$, evaluated in $|x| \geqq 1$. Occasionally, instead of working with homogeneous functions of degree zero we shall work with their restrictions to the unit sphere Σ.

Let $g(x)$ be a function on Σ, and suppose that for each x_0 there is a

neighborhood N_{x_0} of x_0 and an operator in $\mathfrak{H}_{x_0}$ whose Fourier transform h_{x_0} (restricted to Σ) coincides with $g(x)$ in N_{x_0}. Let N_{x_i}, $i=1,2,\cdots$ be a finite collection of such neighborhoods covering Σ. Let further $k_i \geqq 0$ be functions in C^∞, each vanishing outside N_{x_i} and such that $\Sigma k_i(x) > 0$. Then

$$k'_i(x) = k_i(x)\,[\sum_j k_j(x)]^{-1}$$

is also in C^∞ and vanishes outside N_{x_i}. Furthermore $\Sigma k'_i(x) = 1$, and consequently

$$g(x) = \Sigma\, g(x) k'_i(x) = \Sigma\, h_{x_i}(x) k'_i(x),$$

since $h_{x_i}(x) = g(x)$ wherever $k'_i(x) \neq 0$. Since $k'_i(x)$ is a restriction to Σ of a homogeneous function of degree zero of C^∞ in $x \neq 0$, which is in turn the Fourier transform of an operator $\mathcal{K}_i$ in $\mathfrak{A}$, the last expression on the right is precisely the restriction to Σ of the Fourier transform of $\Sigma\, \mathfrak{H}_{x_i}\mathcal{K}_i$, and Theorem 3 is thus established.

To prove Theorem 4 we begin by observing that, as an easy computation shows, if $\alpha > |h(x)|$ and $h(x)$ is in C^∞, then $\beta(h^k) = O(\alpha^k)$ as $k \to \infty$. Consequently it follows from (2.8) that $\| h^k \|_p = O(\alpha^k)$.

We now extend this result to the Fourier transform $h(x)$ of an arbitrary operator $\mathfrak{H}$ in $\mathfrak{A}_p$. Given such an $h(x)$ and $\alpha > |h(x)|$, we take α_0 so that $\alpha > \alpha_0 > |h(x)|$, and $h_0(x)$ in C^∞ so that $A_p \| h - h_0 \|_p + \alpha_0 < \alpha$ and $|h_0(x)| < \alpha_0$. Then, by (1.6),

$$\| h^k \|_p = \| [(h - h_0) + h_0]^k \|_p \leqq \sum_{i=0}^{k} \binom{k}{i} \| (h - h_0)^i h_0^{k-i} \|_p$$
$$= O[\sum_{i=0}^{k} A_p{}^i \binom{k}{i} \| h - h_0 \|_p{}^i \alpha_0{}^{k-i}] = O[(A_p \| h - h_0 \|_p + \alpha_0)^k] = O(\alpha^k).$$

Let now $F(z) = \Sigma\, a_k (z - z_0)^k$ be analytic in $|z - z_0| < 2\epsilon$, $h(x)$ the Fourier transform of an operator in $\mathfrak{A}_p$, and $h(x_0) = z_0$. If we show that $g(x) = F[h(x)]$ coincides with the Fourier transform of an operator in $\mathfrak{A}_p$ in some neighborhood of x_0, a repeated application of this result and Theorem 3 will yield Theorem 4. For this purpose we take $0 \leqq k(x) \leqq 1$ homogeneous of degree zero, equal to 1 in a neighborhood of x_0 and vanishing wherever $|h(x) - h(x_0)| \geqq \epsilon$, and define

$$h'(x) = h(x_0) + k(x)\,[h(x) - h(x_0)].$$

Clearly we have $|h'(x) - h(x_0)| < \epsilon$ and $h'(x) = h(x)$ in a neighborhood of x_0. The series

$$F[h'(x)] = \Sigma\, a_k [h'(x) - h(x_0)]^k$$

coincides with $F[h(x)]$ in a neighborhood of x_0. But since

$$\| [h'(x) - h(x_0)]^k \|_p = O(\epsilon^k),$$

the corresponding series of operators converges in $\mathfrak{A}_p$, and the Fourier transform of its sum is precisely $F[h'(x)]$. Theorem 4 is thus established.

Regarding the Corollary to Theorem 4 we observe that if $h(x)$ does not vanish then $h^{-1}(x)$ satisfies the conditions of Theorem 4 and consequently it is the Fourier transform of an operator in $\mathfrak{A}_p$.

To determine the structure of the maximal ideals in $\mathfrak{A}_p$ we observe that if $h_i = \mathfrak{F}(\mathfrak{H}_i)$ and the $\mathfrak{H}_i$ belong to a proper ideal I in $\mathfrak{A}_p$, the $h_i(x)$ must be necessity have a common zero. For otherwise there would exist a finite number of such $h_i(x)$ without common zero, and the function $h(x) = \Sigma h_i \bar{h}_i > 0$ would be the Fourier transform of an invertible operator in I, and I would not be a proper ideal. Consequently a maximal ideal in $\mathfrak{A}_p$ consists of all operators whose Fourier transform vanish at a point of Σ, and conversely.

MASSACHUSETTS INSTITUTE OF TECHNOLOGY AND
THE UNIVERSITY OF CHICAGO.

REFERENCES.

[1] S. Bochner and K. Chandrasekharan, "Fourier transforms," *Annals of Mathematics Studies*, 19 (1949).

[2] A. P. Calderón and A. Zygmund, "On the existence of certain singular integrals," *Acta Mathematica*, vol. 88 (1952), 85-139.

[3] ———, "On singular integrals," *American Journal of Mathematics*, vol. 78 (1956), pp. 289-309.

ON THE VALUATIONS CENTERED IN A LOCAL DOMAIN.* [1]

By Shreeram Abhyankar.

It is well known that if v is a real zero dimensional valuation of an algebraic function field K/k of n variables and if r is the rational rank of v then the following is true: $r \leqq n$ and if $r = n$ then the value group of v a direct sum of n cyclic groups; see Theorem 1 of [8]. In Section 1, we prove a generalization of this theorem to arbitrary valuations (real or not) centered in an abstract local domain. Namely, let R be a local domain of dimension n with maximal ideal M and quotient field K. Let v be an arbitrary valuation of K having center M in R. Let ρ be the rank and r the rational rank of v, and let d be the transcendence degree of the residue field D of v over R/M. Then we have the following: (1) $d + r \leqq n$, (2) if $d + r = n$ then the value group of v is a direct sum of r cyclic groups and $D/(R/M)$ is finitely generated, (3) if $d + \rho = n$ then v is discrete and $D/(R/M)$ is finitely generated.

In Section 2, we generalize some theorems concerning quadratic sequences of quotient rings on a nonsingular algebraic surface to abstract regular local domains. In Theorem 2, we prove that if f is a nonzero element in a two dimensional regular local domain (R, M) with quotient field K and if v is a valuation of K having center M in R then there exists a quadratic transform (R^*, M^*) of R along v and a basis (x^*, y^*) of M^* such that $f = x^{*a}y^{*b}d$ where a and b are nonnegative integers and d is a unit in R^*, (if v is nonreal, then we assume that R is either algebraic or absolute; for definitions see the beginning of Section 1). The special case of Theorem 2 when R is the quotient ring of a point on an algebraic surface and when v is real plays an important role in the proof of the local uniformization theorem (see Proposition 3 of [2] and Lemma 11.2 of [11]). In Theorem 3 we generalize Zariski's factorization theorem on birational transformations between algebraic surfaces to abstract regular two dimensional local domains. As an incidental remark, we show in Proposition 3 that if (R, M) is a regular n dimensional local domain and if w is a valuation of the quotient field of R with center M in R and of R-dimension $n-1$ then R_w/M_w is a purely transcendental exten-

* Received August 1, 1955.

[1] This work was supported by a research project at Harvard University, sponsored by the National Science Foundation.

sion of a finitely generated extension of R/M. It is a consequence of this proposition that if w is a prime divisor of the second kind having center at a simple subvariety of an algebraic variety then the residue field of w is the function field of a ruled variety. We hope to use the results of this paper in the problem of local uniformization on *absolute* surfaces which we are planning to study.

1. Classification of valuations. We start with some remarks about ordered abelian groups. Let G be an ordered abelian group. Let us recall that the rational rank of G is defined to be the maximum cardinal number (finite or infinite) of a subset H of G such that any finite number $h_1, h_2, \cdots, h_n$ of distinct elements of H are rationally independent, i.e., if $m_1h_1 + m_2h_2 + \cdots + m_nh_n = 0$ where the m_i are integers then $m_1 = m_2 = \cdots = m_n = 0$. We observe that if G has finite rational rank r then we can find r elements $g_1, g_2, \cdots, g_r$ in G such that for any element g in G there exist integers $m_1, m_2, \cdots, m_r$ and a nonzero integer m for which $mg = m_1g_1 + m_2g_2 + \cdots + m_rg_r$ (we call $\{g_1, g_2, \cdots, g_r\}$ a rational basis of G); and conversely if there exist r elements $g_1, g_2, \cdots, g_r$ in G with this property then G is of finite rational rank $n \leqq r$. We assert that the following two conditions are equivalent: (1) G is of finite rational rank r. (2) G is of finite rank ρ and if $(0) = G_0 < G_1 < G_2 < \cdots < G_\rho = G$ is the sequence of isolated subgroups of G then G_i/G_{i-1} is of finite rational rank r_i for $i = 1, 2, \cdots, \rho$. Furthermore, when one and hence both of the above conditions are satisfied, we have the equation: $r = r_1 + r_2 + \cdots + r_\rho$. This is easily verified by writing the elements of G as lexicographically ordered ρ-tuples of real numbers[2] (roughly speaking, if we put together rational bases of G_1/G_0, $G_2/G_1, \cdots, G_\rho/G_{\rho-1}$ then we get a rational basis of G). Given an additive group F of real numbers, we shall say that F is an *integral direct sum*, if F is of finite rational rank a and if we can find elements $f_1, f_2, \cdots, f_a$ in F such that any element f in F can be expressed as: $f = m_1f_1 + m_2f_2 + \cdots + m_af_a$ where the m_i are integers. Again, given an ordered abelian group G, we shall say that G is an integral direct sum if G is of finite rational rank r and if G_i/G_{i-1} is an integral direct sum for $i = 1, 2, \cdots, \rho$ where $(0) = G_0 < G_1 < G_2 < \cdots < G_\rho = G$ is the sequence of isolated subgroups of G. One can easily show that G is an integral direct sum if and only if G is of finite rational rank r and if G contains elements $g_1, g_2, \cdots, g_r$ such that any element g of G can be expressed as: $g = m_1g_1 + m_2g_2 + \cdots + m_rg_r$ where the m_i

[2] See the Appendix.

are integers ($g_1, g_2, \cdots, g_r$ are then necessarily rationally independent; we call $\{g_1, g_2, \cdots, g_r\}$ an integral basis of G). We observe that *G is discrete if and only if G is an integral direct sum and the rational rank of G equals the rank of G.* Finally, note that if H is a subgroup of G of finite index then the rank (respectively, the rational rank) of H equals the rank (respectively, the rational rank) of G and H is an integral direct sum (respectively, discrete) if and only if G is an integral direct sum (respectively, discrete).

We shall consistently use, in this paper, the following *notations.* Let v be a valuation of a field K. We shall denote by R_v the valuation ring of v and by M_v the maximal ideal in R_v. By the rank (respectively, rational rank) of v we shall mean the rank (respectively, rational rank) of the value group of v, also we shall say that v is an integral direct sum if the value group of v is an integral direct sum, and so on. Unless otherwise stated, we shall exclude trivial valuations. When we do want to talk of a trivial valuation, we shall assign to it the zero rank and the zero rational rank, since its value group consists of the zero element alone, and for a trivial valuation we shall consider the designations "integral direct sum" and "discrete" as trivially valid. If v^* is a valuation of a field K^* and v its restriction to a subfield K, then by the v-dimension of v^* is meant the transcendence degree of R_{v^*}/M_{v^*} over R_v/M_v; if v is trivial (over K) then we have that the v-dimension of v^* equals the K-dimension of v, i.e., the transcendence degree of R_{v^*}/M_{v^*} over K. If R is a local ring and M its maximal ideal, we shall indicate this by saying that (R, M) is a local ring. If a local ring R is the quotient ring of an irreducible subvariety of an algebraic variety, then we shall say that R is *algebraic.* By Q and Z we shall denote the field of rational numbers and the domain of ordinary integers respectively. If a local ring R is the quotient ring of a domain A finitely generated over Z (i.e., $A = Z[x_1, x_2, \cdots, x_n]$ with x_i in A) with respect to a prime ideal in A, then we shall say that R is *absolute.* If (R, M) is a local domain and v a valuation of the quotient field of R having center in R (i.e., such that $R_v \supset R$ and $R \cap M_v = M$), then by the R-dimension of v is meant the transcendence degree of R_v/M_v over R/M. Recall that if R is a (commutative) ring and P a nonzero prime ideal in R, then by the dimension of P is meant the maximum value of n such that there exists a strictly ascending chain $P = P_0 < P_1 < P_2 < \cdots < P_n < R$ of prime ideals P_i, and by the rank of P is meant the maximum value of n such that there exists a strictly descending chain $P = P_0 > P_1 > P_2 > \cdots > P_n \supset (0)$ of prime ideals P_i (where P_n is allowed to be the zero ideal if this be prime). Finally, a given integral domain F will be called normal if F is integrally closed in its quotient field.

LEMMA 1. *Let K be a field and K^* an extension of K of finite transcendence degree s. Let v^* be a valuation of K^* and let v be the K-restriction of v^* where we allow v to be trivial. Let d be the v-dimension of v^*. Let r and r^* be the rational ranks of v and v^* and let ρ and ρ^* be the ranks of v and v^* respectively. Then we have the following*: (1) *If r^* is finite, then $r^* + d \leqq r + s$.* (2) *If v is an integral direct sum, if K^*/K is finitely generated, and if $r^* + d = r + s$, then v^* is an integral direct sum and R_{v^*}/M_{v^*} is finitely generated over R_v/M_v (note that $R_v/M_v = K$ if v is trivial).* (1*) *If ρ is finite, then $\rho^* + d \leqq \rho + s$.* (3) *If v is discrete, if K^*/K is finitely generated, and if $\rho^* + d = \rho + s$, then v^* is discrete and R_{v^*}/M_{v^*} is finitely generated over R_v/M_v.*

Proof. First assume that r is finite. We begin by proving the weaker inequality:

(A) $$r^* \leqq r + s.$$

Suppose $s = 0$. Given $0 \neq u \,\varepsilon\, K^*$, let $f(X) = A_0X^n + a_1X^{n-1} + \cdots + a_n$, $a_0 = 1$, be the minimal monic polynomial of u over K. Since $f(u) = 0$, there exist distinct integers i and j such that $v^*(a_iu^{n-i}) = v^*(a_ju^{n-j}) \neq \infty$ and hence $v^*(u) = v(a_i/a_j)/(i-j)$, i.e., the value of u depends rationally on the value of $(a_i/a_j) \,\varepsilon\, K$. Therefore $r^* = r = r + s$. Now suppose $s > 0$ and assume that (A) is true for $s-1$. Let $z_1, z_2, \cdots, z_{s-1}$ be part of a transcendence basis of K^*/K. Let $K_1 = K(z_1, z_2, \cdots, z_{s-1})$, let v_1 be the restriction of v^* to K_1 (v_1 may be trivial), and let r_1 be the rational rank of v_1. By our induction hypothesis, $r_1 \leqq r + s - 1$. If the value of every nonzero element of K^* is rationally dependent on the values of elements of K_1, then $r^* = r_1 \leqq r + s - 1 \leqq r - s$, and we are through. Now suppose that there is a nonzero element z in K^* such that $h = v^*(z)$ does not depend rationally on the values of elements of K_1. Then, by the $s = 0$ case, z is transcendental over K_1. Let $f(X) = f_0 + f_1X + \cdots + f_nX^n$ and $g(X) = g_0 + g_1X + \cdots + g_{n^*}X^{n^*}$ be nonzero elements of $K_1[X]$. Let $a_i = v^*(f_i)$ if $f_i \neq 0$ and $b_i = v^*(g_i)$ if $g_i \neq 0$. Since h depends rationally neither on the a_i nor on the b_i, there exist integers p and q such that $\infty \neq v^*(f_pz^p) < v^*(f_iz^i)$ whenever $i \neq p$ and $f_i \neq 0$, and $\infty \neq v^*(g_qz^q) < v^*(g_iz^i)$ whenever $i \neq q$ and $g_i \neq 0$; i.e., $v^*(f(z)/g(z)) = v^*(f_p/g_q) + (p-q)h$. Thus, *the value of any nonzero element of $K_1(z)$ is of the form $a + mh$ where a is in the value group of v_1 and m is an integer*, i.e., if r_2 is the rational rank of the restriction of v^* to $K_1(z)$ (this restriction may be trivial), then

$$r_2 = r_1 + 1 \leqq r + (s-1) + 1 = r + s.$$

Since $K^*/K_1(z)$ is an algebraic extension, by the case $s=0$, we have $r^*=r_2\leqq r+s$. Thus the induction is complete and (A) has been proved. *Also observe that if v_2 is the restriction of v^* to $K_2=K_1(z)$, then the residue fields of v_1 and v_2 coincide.* For, in the above notation, since $v_2(f_iz^i)>v_2(f_pz^p)$ whenever $i\neq p$ and $f_i\neq 0$, we must have that $f(z)/(f_pz^p)$ belongs to R_{v_2} and that

$$f(z)/(f_pz^p)=1+\sum_{i\neq p}(f_i/f_p)z^{i-p}\equiv 1 \pmod{M_{v_2}}.$$

Similarly, $g(z)/(g_qz^q)$ belongs to R_{v_2} and $g(z)/(g_qz^q)\equiv 1 \pmod{M_{v_2}}$. Now assume that $f(z)/g(z)$ belongs to R_{v_2}. We want to show that we can find e in R_{v_1} with $f(z)/g(z)\equiv e \pmod{M_{v_2}}$. If $f(z)/g(z)$ belongs to M_{v_2}, we can take $e=0$. Now suppose that $f(z)/g(z)$ does not belong to M_{v_2}, i.e., that $v_2(f(z)/g(z))=0$. Since $v_2(f(z)/g(z))=v_2(f_p/_q)+(p-q)h$ and since h does not depend rationally on $v_2(f_p/g_q)$, we must have $p-q=0$, i.e., that $p=q$ and $v_2(f_p/g_p)=0$. Let $e=f_p/g_p$. Then

$$f(z)/g(z)=(f_p/g_p)(f(z)/f_pz^p)(g_pz^p/g(z))\equiv e \pmod{M_{v_2}}$$

since $f(z)/f_pz^p$ and $g_pz^p/g(z)$ are both congruent to one modulo M_{v_2}. This proves our second italicized assertion.

To prove (1) let us retain our assumption that r is finite, and let $\bar{y}_1,\bar{y}_2,\cdots,\bar{y}_d$ be a transcendence basis of R_{v^*}/M_{v^*} over R_v/M_v and fix y_i in R_{v^*} belonging to the residue class $\bar{y}_i$. Let $K'=K(y_1,y_2,\cdots,y_d)$ and let v' be the restriction of v^* to K'. Given $0\neq f(X_1,X_2,\cdots,X_d)$ in $K[X_1,X_2,\cdots,X_d]$ choose a coefficient q of f having minimum v-value and let $F(X_1,X_2,\cdots,X_d)$ $=(1/q)\ f(X_1,X_2,\cdots,X_d)$. Then all the coefficients of $F(X_1,X_2,\cdots,X_d)$ belong to R_v and at least one of them is equal to 1. Let $\bar{F}(X_1,X_2,\cdots,X_d)$ be the polynomial gotten by reducing the coefficients of $F(X_1,X_2,\cdots,X_d)$ modulo M_v. Since $\bar{F}(X_1,X_2,\cdots,X_d)$ has a coefficient equal to 1 and since $\bar{y}_1,\bar{y}_2,\cdots,\bar{y}_d$ are algebraically independent over R_v/M_v, we must have $\bar{F}(\bar{y}_1,\bar{y}_2,\cdots,\bar{y}_d)\neq 0$, i.e., $v^*(F(y_1,y_2,\cdots,y_d))=0$, i.e., $v^*(f(y_1,y_2,\cdots,y_d))$ $=v(q)\neq\infty$, and hence $f(y_1,y_2,\cdots,y_d)\neq 0$. Thus $y_1,y_2,\cdots,y_d$ are algebraically independent over K and *the value groups of v and v' are identical.* Since the transcendence degree of K^*/K' is $s-d$, (1) follows by applying (A) to K^*/K'. Now let $g(X_1,X_2,\cdots,X_d)$ and $h(X_1,X_2,\cdots,X_d)$ be arbitrary nonzero elements of $K[X_1,X_2,\cdots,X_d]$ and let

$$y=f(y_1,y_2,\cdots,y_d)/g(y_1,y_2,\cdots,y_d).$$

Fix coefficients a and b of g and h respectively having minimum v'-values

and let $p=a/b$. Let $G(X_1,X_2,\cdots,X_d)=(1/a)g(X_1,X_2,\cdots,X_d)$ and $H(X_1,X_2,\cdots,X_d)=(1/b)h(X_1,X_2,\cdots,X_d)$. Then, as above,

$$v'(g(y_1,y_2,\cdots,y_d)/h(y_1,y_2,\cdots,y_d))=v'(a/b).$$

Hence $y=g(y_1,y_2,\cdots,y_d)/h(y_1,y_2,\cdots,y_d)$ belongs to $R_{v'}$ if and only if $p=a/b$ belongs to $R_{v'}\cap K=R_v$. Now assume that y does belong to $R_{v'}$. Let $\bar{y}$ and $\bar{p}$ be the residue classes modulo $M_{v'}$ containing y and p respectively. Let $\bar{G}(X_1,X_2,\cdots,X_d)$ and $\bar{H}(X_1,X_2,\cdots,X_d)$ be the polynomials obtained respectively from $G(X_1,X_2,\cdots,X_d)$ and $H(X_1,X_2,\cdots,X_d)$ by reducing their coefficients modulo $M_{v'}$. Since $\bar{H}$ has a coefficient equal to one and since $\bar{y}_1,\bar{y}_2,\cdots,\bar{y}_d$ are algebraically independent over R_v/M_v, we have that

$$\bar{y}=\bar{p}\bar{G}(\bar{y}_1,\bar{y}_2,\cdots,\bar{y}_d)/\bar{H}(\bar{y}_1,\bar{y}_2,\cdots,\bar{y}_d).$$

Therefore $R_{v'}/M_{v'}=(R_v/M_v)(\bar{y}_1,\bar{y}_2,\cdots,\bar{y}_d)$, *and hence in particular* $R_{v'}/M_{v'}$ *is finitely generated over* R_v/M_v.

Now assume that v is an integral direct sum, that K^*/K is finitely generated, and that $r^*+d=r+s$. Let K' and v' be as above. Then v and v' have the same value groups, K^*/K' is a finitely generated extension of transcendence degree $e=s-d=r^*-r$, and $R_{v'}/M_{v'}$ is finitely generated over R_v/M_v. Fix an integral basis $t_1,t_2,\cdots,t_r$ of the value group of v'. Let $x_1,x_2,\cdots,x_e$ be a transcendence basis of K^*/K'. Let $K_i'=K'(x_1,x_2,\cdots,x_i)$, $v_i'=$ the restriction of v^* to K_i', and $r_i'=$ the rational rank of v_i'. Since $r^*=r+e$, we must have, in view of (1), $r_i=r_{i-1}+1$ for $i=1,2,\cdots,e$. Let $v^*(x_i)=t_{r+i}$. By applying the first of the above italicized remarks successively to the extensions $K_1'/K', K_1',\cdots,K_e'/K_{e-1}'$, we conclude that for any nonzero element x of K_e' we have

$$v^*(x)=a+m_{r+1}t_{r+1}+m_{r+2}t_{r+2}+\cdots+m_{r^*}t_{r^*}$$

where a is the value of an element of K' and where $m_{r+1},m_{r+2},\cdots,m_{r^*}$ are integers; since $a=m_1t_1+m_2t_2+\cdots+m_rt_r$ where $m_1,m_2,\cdots,m_r$ are integers, we finally have: $v^*(x)=m_1t_1+m_2t_2+\cdots+m_{r^*}t_{r^*}$. Therefore v_e' is an integral direct sum. Since K^*/K_e' is a finite algebraic extension, the value group of v_e' is a subgroup of the value group of v^* of finite index and hence v^* is an integral direct sum. Now by the second italicized remark above, the residue field of v_e' coincides with the residue field of v'. Since the residue field of v' is finitely generated over the residue field of v and since K^*/K_e' is a finite algebraic extension, we conclude that R_{v^*}/M_{v^*} is finitely generated over R_v/M_v. This proves (2).

The proof of (1*) is entirely similar to that of (1). Finally, assume

that v is discrete, $\rho^* + d = \rho + s$, and that K^*/K is finitely generated. The discreteness of v implies that $\rho = r$. Since by (1), $r^* + d \leqq r + s$ and since $r^* \geqq \rho^*$, it follows that $r^* = \rho^*$ and that $r^* + d = r + s$. Hence by (2), v^* is an integral direct sum and R_{v^*}/M_{v^*} is finitely generated over R_v/M_v. Since $r^* = \rho^*$, v^* is discrete. This proves (3).

COROLLARY 1. *In the notation of the above lemma, assume that v is trivial, i.e., that v^* is a valuation of K^*/K. Then:* (1) $\rho^* + d \leqq r^* + d \leqq s$. *Furthermore, if K^*/K is finitely generated (i.e., if K^*/K is an algebraic function field of dimension s), then we have the following:* (2) *If $r^* + d = s$ then v^* is an integral direct sum and R_{v^*}/M_{v^*} is finitely generated over K.* (3) *if $\rho^* + d = s$ then v^* is discrete and R_{r^*}/M_{r^*} is finitely generated over K.* (4) *If $d = s - 1$ then v^* is real discrete and R_{v^*}/M_{v^*} is finitely generated over K.*

PROPOSITION 1. *Let K be a field which is of finite transcendence degree over its prime field P. Let n be the absolute dimension*[3] *of K and let v be a valuation of K. Let r be the rational rank of v, let ρ be the rank of v, and let d be the absolute dimension of R_v/M_v. Then:* (1) $r + d \leqq n$. *Furthermore, if K is finitely generated over P then we have the following:* (2) *If $r + d = n$ then v is an integral direct sum and R_v/M_v is finitely generated over its prime field.* (3) *If $\rho + d = n$ then v is discrete and R_v/M_v is finitely generated over its prime field.* (4) *If $d = n - 1$ then v is real discrete and R_v/M_v is finitely generated over its prime field.*

Proof. Let w be the restriction of v to P where w may be trivial. If P is of nonzero characteristic then w is trivial, $n =$ the transcendence degree of K over P, and $R_w/M_w = P$; and hence the proposition follows from the previous corollary. Now assume that P is of characteristic zero, i.e., that $P = Q$ (the field of rational numbers). Let D be the w-dimension of v, d^* the absolute dimension of R_w/M_w, and r^* the rational rank of w. If w is trivial, then $r^* = 0$ and $R_w/M_w = Q$, i.e., $d^* = 1$ and hence $r^* + d^* = 1$. If w is nontrivial, then w must be the real discrete valuation given by a prime number p of Z and hence $r^* = 1$ and $R_w/M_w = Z/(pZ)$, i.e., $d^* = 0$ and hence $r^* + d^* = 1$. Thus in both the cases, w is discrete, $r^* + d^* = 1$, and R_w/M_w is its own prime field. Hence $r + D \leqq r^* + (n - 1)$ if and only if $r + (d - d^*) \leqq r^* + (n - 1)$, i.e., if and only if

$$r + d \leqq (d^* + r^*) + (n - 1) = 1 + (n - 1) = n.$$

[3] Let D be an integral domain with quotient field K such that K is of finite transcendence degree s over its prime subfield. Then we define the absolute dimension of D to be s or $s + 1$ according as K is of nonzero or zero characteristic respectively.

Thus $r+D \leqq r^* + (n-1)$ if and only if $r+d \leqq n$ and $r+D = r^* + (n-1)$ if and only if $r+d=n$. Hence (1), (2), (3) follow respectively from parts (1), (2), (3) of Lemma 1 applied to the extension K/Q. Finally, (4) follows from (1) and (3).

DEFINITION 1. *In the notation of Proposition 1, if $d=n-1$ then we shall say that v is a prime divisor of K.*

Remark 1. Let K be a field of absolute dimension two which is finitely generated over its prime field P of characteristic p; we can express this by saying that *K is the function field of an absolute surface of characteristic p.* Let v be a valuation of K. It follows by Proposition 1 that v is of one of the following four types: (1) v is a prime divisor of K, i.e., v is real discrete and R_v/M_v is of absolute dimension one. If $p \neq 0$ then R_v/M_v is an algebraic function field of one variable over P. If $p=0$ then R_v/M_v is either a finite algebraic extension of Q or it is an algebraic function field of dimension one over a prime field of nonzero characteristic. In either case, R_v/M_v is the function field of an *absolute curve.* (2) v is discrete of rank two. (3) v is real of rational rank one, but v is not a divisor of K. (4) v is real of rational rank two; in this case v is necessarily an integral direct sum. In cases (2), (3), and (4), R_v/M_v is an algebraic extension of a prime field of characteristic $q \neq 0$ and hence R_v/M_v is perfect; if $p \neq 0$ then $q=p$; in cases (2) and (4), R_v/M_v, being finitely generated over its prime field, is necessarily finite.

If $p \neq 0$ then K/P is an algebraic function field of two variables, i.e., an absolute surface (respectively, an absolute n-dimensional variety) of characteristic $p \neq 0$ is simply an algebraic surface (respectively, an algebraic n-dimensional variety) over the prime field of characteristic p; in this case the above classification of valuations is well known. The present remark signifies that a parallel situation prevails for absolute surfaces of characteristic zero, i.e., algebraic curves over Q considered as surfaces.

PROPOSITION 2. *Let (R,M) be a local domain of dimension n with quotient field K. Let v be a real valuation of K with center M in R. Let d be the R-dimension of v and let r be the rational rank of v. Then $d+r \leqq n$. Furthermore, if $d+r=n$ then v is an integral direct sum and R_v/M_v is finitely generated over R/M.*

We shall divide the proof of this proposition into several lemmas.

LEMMA 2. *Proposition 2 is true in case R is the power series ring $k[[x_1, x_2, \cdots, x_n]]$ in n variables $x_1, x_2, \cdots, x_n$ over a field k.*

Proof. Let

$$K^* = k(x_1, x_2, \cdots, x_n), \quad A = k[x_1, x_2, \cdots, x_n], \quad P = (x_1, x_2, \cdots, x_n)A,$$

$R^* = A_P$, $M^* = PR^*$, and $v^* =$ the restriction of v to K^*. Then (R, M) is the completion of (R^*, M^*) and hence by Lemma 12 of [2], v^* has center M^* in R^* and v^* has the same residue field and value group as v. Since $R/M = R^*/M^*$, d is also the r^*-dimension of v^*. Since $R^*/M^* = k$, the present lemma follows by applying Lemma 1 to the extension K^*/k.

LEMMA 3. *Proposition 2 is true in case R is the power series ring $R_w[[x_1, x_2, \cdots, x_{n-1}]]$ in $n-1$ variables $x_1, x_2, \cdots, x_{n-1}$ over a complete valuation ring R_w of a real discrete valuation w such that R_w is of characteristic zero and $M_w = pR_w$ where p is an ordinary prime number.*

Proof. Let k be the quotient field of R_w, $K^* = k(x_1, x_2, \cdots, x_{n-1})$, $A = R_w[x_1, x_2, \cdots, x_{n-1}]$, $P = A \cap M$, $R^* = A_P$, $M^* = PR^*$, and $v^* =$ the K^*-restriction of v. Then it is easily verified that $M^* = (x_1, x_2, \cdots, x_{n-1})R^*$, that (R^*, M^*) is a regular local domain of dimension n, and that (R, M) is the completion of (R^*, M^*); [Proof: $M = (x_1, x_2, \cdots, x_{n-1}, p)R$, $M^i \cap R^* = M^{*i}$ and hence R and R^* are concordant, and so on]. By the argument used in the proof of Lemma 12 of [2], it follows that v^* has center M^* in R^* and v^* has the same value group and residue field as v. Since $R/M = R^*/M^*$, d must also be the R^*-dimension of v^*. Since $M_{v^*} \cap R_w = M^* \cap R_w = M_w$ and $R^*/M^* = R_w/M_w$ (Lemma 5 of [3]), since $R_w \subset R_{v^*}$ and since R_w is a maximal subring of k, it follows that w is the k-restriction of v^* and that d is also the w-dimension of v^*. The present lemma now follows by applying Lemma 1 to the extension K^*/k.

LEMMA 4. *Proposition 2 is true if R is an unramified complete regular local ring.*

Proof. By Theorem 15 of [3], R is isomorphic either to the ring described in Lemma 2 or to the ring described in Lemma 3. Therefore, Lemma 4 follows from Lemmas 2 and 3.

LEMMA 5. *Proposition 2 is true if R is complete.*

Proof. By Theorem 16 of [3], R is a finite module over a subring S which is an unramified complete regular local ring having the same dimension and residue field as R. Hence Lemma 5 follows from Lemma 4 by observing that valuations do not change their character in passing to a finite algebraic extension.

LEMMA 6. *Proposition 2 is true.*

Proof. Since v is real, we can employ the technique used in the proof of Theorem 1 of [13] to pass from the case of complete local domains to the general case; Lemma 6 then follows from Lemma 5.

THEOREM 1. *Let (R, M) be a local domain of dimension n with quotient field K and let v be a valuation of K with center M in R. Let d, r, and ρ be respectively the R-dimension, the rational rank, and the rank of v. Then: (1) $d+r \leqq n$. (2) If $d+r=n$ then v is an integral direct sum and R_v/M_v is finitely generated over R/M. (3) If $d+\rho=n$ then v is discrete and R_v/M_v is finitely generated over R/M. (4) If $d=n-1$ then v is real discrete and R_v/M_v is finitely generated over R/M.*

Proof. We observe that (3) and (4) follow at once from (1) and (2), and we proceed to prove (1) and (2) by applying induction to n. Suppose first that $n=1$. Let D be the integral closure of R in K. Then D is a Dedekind domain (§ 39 of [5]) and hence we must have $R_v \supset D$ and $M_v \cap D$ $=$ a minimal prime idea H in D, and hence $D_H \subset R_v$. Since D_H is a real discrete valuation ring, it is a maximal subring of K and hence $D_H = R_v$, i. e., v is real discrete. Hence, for $n=1$, (1) and (2) follow from Proposition 2. Now let $n>1$ and assume that (1) and (2) are true for all smaller values of n. If v is real then (1) and (2) follow from Proposition 2. So asume that v is nonreal. Given a nonzero prime ideal E in R_v, $(R_v)_E$ is the valuation ring of a valuation of K (see [4]), and hence $E \cap R$ must be a nonzero prime ideal in R (since K is the quotient field of R). Since the set T of nonzero prime ideals in R is simply ordered by inclusion [4], the set T^* of the R-contractions of the members of T is also simply ordered by inclusion. Since any chain of prime ideal in R is of length at most n, T^* must be a finite set. Therefore $P = \bigcap_{F \varepsilon T^*} F$ is *the* smallest ideal in T^*, and hence P is a nonzero prime ideal. Let $B = \bigcap_{E \varepsilon T} E$. Then $B \cap R = P \neq (0)$ and hence $B \neq (0)$. Therefore $(R_v)_B$ is the valuation ring of a nontrivial valuation w of K. Since B is the minimal nonzero prime ideal of R_v, w must be real. Let $K^* = R_w/M_w$, $\bar{R} = R/P$, $\bar{M} = M/P$, $\bar{K} =$ the quotient field of $\bar{R}$ in K^*, $v^* =$ the valuation of K^* induced by v (v^* is nontrivial), $\bar{v} =$ the restriction of v^* to $\bar{K}$ where $\bar{v}$ may be trivial. Let $S = R_P$ and $N = PS$. Then w has center N in S. Let d' be the S-dimension of w. Then $d' =$ the transcendence degree of $K^*/\bar{K}$, since $\bar{K} = S/N$. Let $\bar{r}$, r^*, and r' be the rational ranks of $\bar{v}$, v^*, and w respectively. We observe that $r = r^* + r'$

and we divide the remaining argument into two cases according as $P = M$ or $P \neq M$.

Case 1, $P = M$. Then $\bar{R} = \bar{K}$, $\bar{M} = (0)$ and $\bar{v}$ is trivial. Also, the $\bar{K}$-dimension of $v^* =$ the R-dimension of $v = d$. Applying Proposition 2 to the real valuation w we get $d' + r' \leqq n$. Applying Corollary 1 of Lemma 1 to the extension $K^*/\bar{K}$ we get $r^* + d \leqq d'$. Therefore $r + d = r^* + r' + d \leqq n$. Now assume that $r + d = n$. Then $d' + r' = n$ and $r^* + d = d'$. Since $d' + r' = n$, it follows by Proposition 2 that w is an integral direct sum and $K^*/\bar{K}$ is finitely generated. Since $r^* + d = d'$ and since $K^*/\bar{K}$ is finitely generated, it follows by Corollary 1 of Lemma 1 that v^* is an integral direct sum and $R_{v^*}/M_{v^*} = R_v/M_v$ is finitely generated over $\bar{K} = R/M$. Since v is composed of v^* and w and since v^* and w are integral direct sums, it follows that v is also an integral direct sum.

Case 2, $P \neq M$. Let $\bar{n}$ and m be the dimensions of the local rings $\bar{R}$ and S respectively. Then $\bar{n} = \dim P$ and $m = \operatorname{rank} P$. Therefore

$$\bar{n} + m \leqq n. \tag{A}$$

Since $m > 0$, we must have $0 < \bar{n} < n$. Now $\bar{v}$ is a nontrivial valuation of $\bar{K}$ having center $\bar{M}$ in $\bar{R}$. Let $\bar{r}$ be the rational rank of $\bar{v}$ and let $\bar{d}$ be the $\bar{R}$-dimension of $\bar{v}$. Let d^* be the $\bar{v}$-dimension of v^*. Then

$$\bar{d} + d^* = d. \tag{B}$$

Since w is a real valuation with center N in S, it follows by Proposition 2 that

$$r' + d' \leqq m. \tag{C}$$

Applying our induction hypothesis to the valuation $\bar{v}$ having center $\bar{M}$ in $\bar{R}$ we get

$$\bar{r} + \bar{d} \leqq \bar{n}. \tag{D}$$

By Lemma 1,

$$r^* + d^* \leqq \bar{r} + d'. \tag{E}$$

Adding the inequalities (C), (D), and (E) we get

$$r' + \bar{r} + r^* + d' + \bar{d} + d^* \leqq m + \bar{n} + \bar{r} + d',$$

i. e.,

$$r' + r^* + \bar{d} + d^* \leqq m + \bar{n}.$$

Since $r' + r^* = r$ and since by (B), $\bar{d} + d^* = d$, we obtain:

$$r + d \leqq m + \bar{n}. \tag{F}$$

Therefore, in view of (A), we conclude that $r+d \leqq n$. Now assume that $r+d=n$. Then by (A) and (F) we must have $r+d=m+\bar{n}$. Since $r=r'+r^*$ and since by (B), $d=\bar{d}+d^*$, we obtain:

$$r'+\bar{r}+r^*+d'+\bar{d}+d^*=m+\bar{n}+\bar{r}+d'.$$

Therefore by (C), (D), and (E), it follows that:

(C′) $$r'+d'=m.$$

(D′) $$\bar{r}+\bar{d}=\bar{n}.$$

(E′) $$r^*+d^*=\bar{r}+d'.$$

By our induction hypothesis, (D′) implies that $\bar{v}$ is an integral direct sum and that $R_{\bar{v}}/M_{\bar{v}}$ is finitely generated over $\bar{R}/\bar{M}=R/M$. Therefore, in view of (E′), it follows by Lemma 1, that v^* is an integral direct sum and that $R_v/M_v=R_{v^*}/M_{v^*}$ is finitely generated over $R_{\bar{v}}/M_{\bar{v}}$ and hence over R/M. Again, in view of (C′), it follows by Proposition 2 that w is an integral direct sum. Since w and v^* are both integral direct sums, we finally conclude that v is also an integral direct sum.

Thus the induction is complete and the theorem has been established.

DEFINITION 2. *In the notation of Theorem 1, if v is of R-dimension $n-1$ then we shall say that v is a prime divisor for (R,M).*

Remark 2. In the notation of Theorem 1, let $n=2$. Then we get a classification of valuations v of K centered in R which is parallel to the one given in Remark 1 for the case of absolute surfaces. Namely, v is of one of the following four types: (1) v is a prime divisor for (R,M), i. e., $d=1$ and v is real discrete. (2) $d=0$ and $\rho=2$; in this case, necessarily, $r=\rho=2$, i. e., v is discrete. (3) $d=0$ and $\rho=r=1$. (4) $d=0$, $\rho=1$ and $r=2$; in this case v is necessarily an integral direct sum. In cases (1), (2), and (4), R_v/M_v is necessarily finitely generated over R/M. Also note that if v is of positive R-dimension, then v is necessarily of type (1), i.e., v is a prime divisor.

2. Quadratic transformations.

LEMMA 7. *Let $R_1 \subset R_2 \subset \cdots$ be a sequence of normal integral domains with a common quotient field K such that R_i contains a unique maximal ideal M_i and $M_{i+1} \cap R_i = M_i$ for $i=1,2,\cdots$. Assume that $\bigcup_{i=1}^{\infty} R_i$ is not the valuation ring of a valuation of K. Then there exist infinitely many valua-*

tions w of K which have center M_i in R_i and for which R_w/M_w is of positive transcendence degree over R_i/M_i for each i.

Proof. The proof is essentially the same as the one given by Zariski in the special case when the (R_i, M_i) are quotient rings of corresponding subvarieties under a sequence of birational transformations, see the theorem on page 25 of [14]. To outline the main idea of the proof, first observe that for an arbitrary commutative ring A with identity the following two statements are equivalent: (1) A contains a unique maximal ideal. (2) The set of nonunits in A is an ideal.

Now let $R = \bigcup_{i=1}^{\infty} R_i$ and $M = \bigcup_{i=1}^{\infty} M_i$. We may canonically assume that $R_1/M_1 \subset R_2/M_2 \subset \cdots$. Let $D = \bigcup_{i=1}^{\infty} R_i/M_i$. Then D is a field and $R/M = D$, i.e., M is a maximal ideal in the domain R. Also observe that R is normal. Since R is not a valuation ring, there exists $x \varepsilon K$ with $x \notin R$ and $(1/x) \notin R$. The canonical homomorphism h of R onto D can be uniquely extended to a homomorphism H of $R[x]$ onto $D[x]$ for which $H(x) = X$, where X is a transcendental over D; see pp. 26-27 of [14]. Let p be any one of the infinitely many prime ideals in $D[X]$. Let $P = H^{-1}(p)$. By the theorem of existence of valuations, there exists a valuation w of K having center P in $R[x]$. Since $R_w/M_w \supset D(x)$, it is clear that w has the required properties. The infinitely many choices of p gives us infinitely many w of the required type.

LEMMA 8. *Let (R, M) be a two dimensional normal local domain with quotient field K. Let P be a minimal prime ideal in R. Then:* (1) *R_P is the valuation ring of a real discrete valuation w of K;* (2) *there exists at least one and at most a finite number of valuations v of K having center M in R which are composed with w, i.e., for which $R_v \subset R_w$; and* (3) *each such valuation v is discrete of rank two and R_v/M_v is a finite algebraic extension of R/M (hence in particular v is of R-dimension zero).*

Proof. (1) is proved on page 103 of [5]. The proof of (2) is the same as in Lemma 11 of [2] and is as follows: $R/(R \cap M_w)$ is a local domain of dimension one with quotient field R_w/M_w. Hence the integral closure of $R/(R \cap M_w)$ in R_w/M_w is a Dedekind domain D with a finite number of prime ideals $P_1, P_2, \cdots, P_h$ (§ 39 of [5]). Let v_i^* be the real discrete valuation of R_w/M_w with $R_{v_i^*} = D_{P_i}$. Let v_i be the valuation of R_w/M_w which is composed of w and v_i^*. Then $v_1, v_2, \cdots, v_h$ are exactly the valuations described in (2). Finally, (3) follows at once from Theorem 1.

LEMMA 9. *Let (R, M) be a two dimensional regular local domain with quotient field K. Then: (1) R is a unique factorization domain, i.e., equivalently, every minimal prime ideal in R is principal; and (2) there exist infinitely many valuations of K having center M in R which have R-dimension zero and which are discrete of rank two.*

Proof. (1) follows by Satz 8 and Satz 9 of Krull [7].[4] To prove (2) it is enough to show, in view of Lemma 8, that there exist infinitely many relatively prime irreducible nonunits in R. To show this, let x, y be a minimal basis of M. Let $\bar{P} = xR$, $\bar{R} = R/P$, $\bar{M} = M/P$, $\bar{K} = R_P/(PR_P)$, $\bar{y} =$ the residue class modulo P containing y, and $w =$ the real discrete valuation of K with $R_w = R_P$. Then $\bar{M} = \bar{y}\bar{R}$ and hence $\bar{R}$ is a regular one dimensional local domain, i.e., $\bar{R} = R_{\bar{v}}$ where v is a real discrete valuation of $\bar{K}$ with $\bar{v}(\bar{y}) = 1$. Let v be the valuation of K which is composed of w and $\bar{v}$, and let us write the elements of the value group of v as lexicographically ordered pairs of integers. Let $x_m = x + y^m$ where m is a positive integer. Since (x_m, y) is a basis of M, x_m is an irreducible nonunit. Since $v(X_m) = (0, \bar{v}(\bar{y}^m)) = (0, m)$, we have that $v(x_m) \neq v(x_n)$ whenever $m \neq n$. Therefore $x_1, x_2, x_3, \cdots$ are infinitely many pairwise relatively prime irreducible nonunits in R.

Now let us recall the notion of a quadratic transform.

LEMMA 10. *Let (R, M) be a regular local domain of dimension $s > 1$ and let $x_1, x_2, \cdots, x_s$ be a minimal basis of M. Let v be a valuation of the quotient field K of R having center M in R. Suppose we have arranged the x_i so that $v(x_1) \leqq v(x_i)$ for $i = 1, 2, \cdots, s$. Let $A = R[x_2/x_1, x_3/x_1, \cdots, x_s/x_1]$, $P = A \cap M_v$, $S = A_P$ and $N = PS$. Then (S, N) is a regular local domain of dimension $t \leqq s$, v has center N in S and if by d and d^* we denote respectively the R-dimension and the S-dimension of v then we have: $s - t = d - d^*$.*

Proof. By Lemma 3 of [1], $A/(x_1A)$ can be canonically identified with a polynomial ring $A^* = (R/M)(X_2, X_3, \cdots, X_s)$ in $s - 1$ variables over R/M. Let $P^* = P/(x_1A)$. Then $A^*_{P^*}$ is a regular local ring of dimension $h \leqq s - 1$. Fix $y_2^*, y_3^*, \cdots, y_{h+1}^*$ in A^* such that

$$P^*A^*_{P^*} = (y_2^*, y_3^*, \cdots, y_{h+1}^*)A^*_{P^*}.$$

[4] We take this opportunity in pointing out that there is obviously a misprint in the statement of satz 9 of Krull [7], namely that the term "Stellenring" should be replaced by "p-reihenring." For in the proof it is used that the Anfang forms are binary which is true only for p-reihenrings; and further, the theorem is certainly not true for arbitrary two dimensional complete local rings.

Fix y_i in A belonging to the residue class y_i^*. Then it is clear that $(x_1, y_2, y_3, \cdots, y_{h+1})S = N$ and that

$$t = \dim S = \operatorname{rank} N = 1 + \operatorname{rank} P^* = 1 + h.$$

Therefore S is regular. Also $t \leqq 1 + h \leqq (s-1) + 1 = s$. Since $N \cap R = M$, we may canonically assume that $R/M \subset S/N \subset R_v/M_v$. Since the transcendence degree of $S/N = A^*/P^*$ over R/M is $s - 1 - h = s - t$, we conclude that: $s - t = (R\text{-dimension of } v) - (S\text{-dimension of } v) = d - d^*$.

DEFINITION 3. *In the notation of the above lemma, S is called "the first (or immediate) quadratic transform of R along v." Let now $R_0 = R$ and let R_i be the first quadratic transform of R_{i-1} along v assuming that* $\dim R_{i-1} > 1$. *If* $\dim R_i > 1$, *for* $i = 1, 2, \cdots, n-1$ *then R_n will be defined and we shall say that "R_n is the n-th quadratic transform of R along v." If S is the n-th quadratic transform of R along v for some n, we shall say that "S is a quadratic transform of R along v." Finally, if S is a quadratic transform of R along some valuation v having center M in R, we shall say that "S is a quadratic transform of R."*

DEFINITION 4. *Let (R, M) and (S, N) be local domains with a common quotient field K. If $S \supset R$ and $N \cap R = M$ then we shall say that (S, N) has center M in R. If (S, N) has center M in R and if there exists a finite set of elements $x_1, x_2, \cdots, x_n$ in S such that $S = A_P$ and $N = PS$ where $A = R[x_1, x_2, \cdots, x_n]$ and $P = A \cap N$ then we shall say that "(S, N) is a finite transform of (R, M)."*

LEMMA 11. *Let (R, M) and (S, N) be local domains with a common quotient field K. Then we have the following:* (1) *If S is a finite transform of R and if R is absolute (respectively, algebraic) then S is absolute (respectively, alegbraic).* (2) *If R is regular and if S is a quadratic transform of R then S is a finite transform of R.*

Proof. Assume that S is a finite transform of R and fix $x_1, x_2, \cdots, x_n$ in S such that $S = A_P$ and $N = PS$ where $A = R[x_1, x_2, \cdots, x_n]$ and $P = A \cap N$. Let k = the ground field or $k = Z$ according as R is algebraic or absolute respectively. Then we can find a finite number of elements $x_{n+1}, x_{n+2}, \cdots, x_m$ in R such that $R = A^*_{P^*}$ where $A^* = k[x_{n+1}, x_{n+2}, \cdots, x_m]$ and $P^* = A^* \cap M$. Let $A' = k[x_1, x_2, \cdots, x_m]$ and $P' = A' \cap N$. Then it is easily seen that $S = A'_{P'}$ and $N = P'S$ and hence S is absolute or algebraic according as R is absolute or algebraic respectively. The proof of (2) is entirely similar.

PROPOSITION 3. *Let (R, M) be an n-dimensional regular local domain with quotient field K with $n > 1$, and let v be a valuation of K which is a prime divisor for R. Then the quadratic sequence along v starting from R is necessarily finite, i.e., if $R = R_0$ and if R_i is the first quadratic transform of R_{i-1} along v provided $\dim R_{i-1} > 1$ then for some integer h we have that R_h is one dimensional; we also have: $R_h = R_v$. Furthermore, there exists a field T with $R/M \subset T \subset R_v/M_v$ such that T is finitely generated over R/M and R_v/M_v is a pure transcendental extension of T of finite positive transcendence degree (we may express this by saying that "R_v/M_v is a ruled extension of R/M").*

Proof. Assume the contrary, i.e., that the quadratic sequence $R = R_0 < R_1 < R_2 < \cdots$ along v is infinite. It then follows, by Lemma 10, that there exists an integer s such that $\dim R_t = \dim R_s$ and R_t-dimension of $v = m - 1$ whenever $t \geqq s$, where we have set $m = \dim R_s$. Let $S = \bigcup_{i=0}^{\infty} R_i$ and $N = \bigcup_{i=0}^{\infty} M_i$ where M_i is the maximal ideal in R_i. Then, as in the proof of Lemma 7, N is the unique maximal ideal in S and $\bigcup_{i=s}^{\infty} R_i/M_i = S/N$. Since, as in the proof of Lemma 10, R_{t+1}/M_{t+1} is an algebraic extension of R_t/M_t whenever $t \geqq s$, it follows that S/N is an algebraic extension of R_t/M_t for any $t \geqq s$. Now v has center M_i in R_i for each i and by Theorem 1, v is real discrete. Hence, if S were not the valuation ring of a valuation of K then, following the considerations of pages 27-28 of [14], we would reach a contradiction. Therefore, S must be the valuation ring of a valuation w of K. Since R_s has a first quadratic transform R_{s+1}, it follows by the definition of quadratic transforms that $m > 1$, i.e., that v is of positive R_t-dimension whenever $t \geqq s$. Now $R_v \supset R_w$ and

$$R_w \cap M_v = S \cap M_v = \left(\bigcup_{i=0}^{\infty} R_i\right) \cap M_v = \left(\bigcup_{i=0}^{\infty} (R_i \cap M_v)\right) = \bigcup_{i=0}^{\infty} M_i = N = M_w.$$

Therefore $v = w$, i.e., $R_v/M_v = S/N$ is an algebraic extension of R_s/M_s. Thus our assumption that the quadratic sequence $R_0 < R_1 < R_2 < \cdots$ is infinite is absurd. Let R_h be the last member of this sequence. Then R_h is a regular one dimensional local domain, i.e., R_h is the valuation ring of a real discrete valuation of K. Since $R_h \subset R_v$, we must have $R_h = R_v$. Now let $T = R_{h-1}/M_{h-1}$ and let d be the dimension of R_{h-1}. Then $d > 1$. Let $x_1, x_2, \cdots, x_d$ be a minimal basis of M_{h-1} arranged so that $v(x_1) \leqq v(x_i)$ for $i = 1, 2, \cdots, d$. Let $A = R_{h-1}[x_2/x_1, x_3/x_1, \cdots, x_d/x_1]$ and $P = A \cap M_h$. As in Lemma 3 of [1], we may identify $A/(x_1A)$ with the polynomial ring

$A^* = T[Y_2, Y_3, \cdots, Y_d]$. Let $P^* = P/(x_1A)$. As in the proof of Lemma 10, $\operatorname{rank} P^* = (\dim R_h) - 1 = 0$, i. e., $P^* = (0)$. Since $R_v/M_v = R_h/M_h$ is isomorphic to A^*/P^*, i. e., to $T(Y_2, Y_3, \cdots, Y_d)$, we conclude that $R_vM/_v$ is a pure transcendental extension of T of transcendence degree $d - 1 > 0$. By Theorem 1, R_v/M_v is finitely generated over R/M and hence T is also finitely generated over R/M; this also follows from the fact that R_{h-1} is a finite transform of R.

Proposition 4. *Let V be an r-dimensional algebraic variety with function field K/k, let v be a prime divisor of K/k, and let W be the center of v on V. Let s be the dimension of W. Assume that v is of the second kind for V (i. e., that $s < r - 1$) and that W is simple for V. Let V^* be any other projective model of K/k such that V is of the first kind for V^* and such that V^* is normal at the center of W^* of v on V^*. Let L/k be the function field of W^*. Then we have the following:* (1) *W^* is a ruled variety, i. e., we can find a field T with $k \subset T \subset L$ such that L is a pure transcendental extension of T of positive transcendent degree.* (2) *If $r = 2$, then we can find a field T with $k \subset T \subset L$ such that T/k is a finite algebraic extension and L/T is a pure transcendental extension of transcendence degree one.* (3) *If $r = 2$ and if k is algebraically closed, then L/k is a pure transcendental extension of transcendence degree one, i. e., W^* is a rational curve.*

Proof. (1) follows from Proposition 3 by observing that $R_v/M_v = L$ and that if (R, M) denotes the quotient ring of W on V then $k \subset R/M \subset R_v/M_v$. Again, (2) follows from (1) and (3) follows from (2).

Lemma 12. *Let $R_0 < R_1 < R_2 < \cdots$ be a strictly ascending sequence of two dimensional regular local domains with a common quotient field K, let M_i be the maximal ideal in R_i, and let $S = \bigcup_{i=0}^{\infty} R_i$. Assume that R_{i+1} is a quadratic transform of R_i for $i = 0, 1, 2, \cdots$. Then:* (1) *S is the valuation ring of a valuation v^* of K such that v^* has center M_i in R_i and such that v^* is of R_i-dimension zero for each i. Furthermore:* (2) *if v is any valuation of K with center M_i in R_i for $i = 0, 1, 2, \cdots$, then $v = v^*$.*

Proof. By introducing all the intermediate successive quadratic transforms between R_i and R_{i+1} for each i, we may assume without loss of generality, since this would not change S, that R_{i+1} is a first quadratic transform of R_i for $i = 0, 1, 2, \cdots$. Now suppose, if possible, that S is not the valuation ring of a valuation of K. Then by Lemma 7, there exists a valuation w of K having center M_i in R_i and of positive R_i-dimension for each i; and hence

by Theorem 1, w is a prime divisor for R_0. Therefore by Proposition 3, the sequence $R_0 < R_1 < R_2 < \cdots$ must be finite, which is absurd. Hence $S = R_{v^*}$ where v^* is a valuation of K. Since, as in the proof of Lemma 7, $M_{v^*} = \bigcup_{i=0}^{\infty} M_i$, it follows that

$$R_i \cap M_{v^*} = R_i \cap (\bigcup_{j=0}^{\infty} M_j) = R_i \cap (\bigcup_{j=i+1}^{\infty} M_j) = \bigcup_{j=i+1}^{\infty} (R_i \cap M_j) = M_i,$$

i.e., v^* has center M_i in R_i for each i. Again by Proposition 3 and Theorem 1, it follows that the R_i-dimension of v^* is zero for each i. Finally, if v is any other valuation of K with center M_i in R_i for each i then $R_v \supset \bigcup_{i=0}^{\infty} R_i = R_{v^*}$ and $M_v \cap R_{v^*} = \bigcup_{i=0}^{\infty} (M_v \cap R_i) = \bigcup_{i=0}^{\infty} M_i = M_{v^*}$ and hence $v = v^*$.

LEMMA 13. *Let (R, M) be a regular two dimensional local domain with quotient field K, let P be a minimal prime ideal in R, let $\bar{R} = R/P$ and $\bar{M} = M/P$. Let $\bar{K}$ be the quotient field of $\bar{R}$ and let T be the integral closure of $\bar{R}$ in $\bar{K}$. Assume that R is the quotient ring of a point either on an algebraic surface or on an absolute surface (i.e., K is a one dimensional algebraic function field over Q and R is the quotient ring of a finitely generated domain over Z with respect to a maximal ideal). Then T is a finite R-module.*

Proof. The case of algebraic surfaces is well known (see p. 511 of [10]) and so we may assume that R is the quotient ring of a point on an absolute surface. Then we can find $z_1, z_2, \cdots, z_s$ in K such that K is the quotient field of $R^* = Z[z_1, z_2, \cdots, z_s]$ and such that $R = R^*_{M^*}$ where M^* is a maximal ideal in R^*. Let $P^* = R^* \cap P$, $\bar{R}^* = R^*/P^*$, $\bar{M}^* = M^*/P^*$, and $\bar{z}_i =$ the P^*-residue class of z. Then $\bar{R} = \bar{R}^*_{\bar{M}^*}$. If $Z \cap P^* \neq (0)$ then $Z \cap P^* = pZ$ where p is a prime number and $\bar{R}^* = (Z/pZ)[\bar{z}_1, z_2, \cdots, \bar{z}_s]$; hence we are again in the algebro-geometric case. Now assume that $Z \cap P^* = (0)$. Let $S = R^*_{P^*}$, $N = P^*S$ and $\bar{S} = S/N$. Then $\bar{S} = \bar{K}$. Also $Q \subset S$ and hence $S = S^*_{N^*}$ where $S^* = Q[z_1, z_2, \cdots, z_s]$ and $N^* = S^* \cap N$. Since S^* is of transcendence degree one over Q, it follows that $\bar{K}$ is a finite algebraic extension of Q. Thus it is enough to prove *the following assertion: Let L be a finite algebraic extension of Q, let A be a proper subdomain of L having L as quotient field, and let B be the integral closure of A in L. Then B is a finite A-module.*

First let us recall that if D is a commutative ring with 1, and W the set of all maximal ideal in D then $D = \bigcap_{w \in W} D_w$. For if $d \in \bigcap_{w \in W} D_w$ then $d = e_w/f_w$ with $e_w, f_w \in D$ and $f_w \notin w$. Now since the ideal E generated by all the f_w

is not contained in any member of W, we must have $E = D$. Therefore we can write $1 = f_{w_1}g_1 + f_{w_2}g_2 + \cdots + f_{w_h}g_h$ with g_i in D and hence $d = e_{w_1}g_1 + e_{w_2}g_2 + \cdots + e_{w_h}g_h \in D$. Therefore $D = \bigcap_{w \in W} D_w$. To prove the italicized assertion, let Y be the integral closure of Z in L. Then Y is a finite Z-module; let $y_1, y_2, \cdots, y_n$ be a module basis of Y over Z. Let $C = A[y_1, y_2, \cdots, y_n]$. Then it is clear that $y_1, y_2, \cdots, y_n$ is a module basis of C over A. Let M be a maximum ideal in C, and let $m = Y \cap M$. Then $Y_m \subset C_M$. If $m = (0)$ then $Y_m = L$ and if $m \neq (0)$ then Y_m is the valuation ring of real discrete valuation of L, i. e., Y_m is a maximal subring of L. Therefore $Y_m = C_M$. Hence by the above observation, C is an intersection of valuation rings. Therefore C is integrally closed in L. Since $C \subset B$, we must have $C = B$, i. e., B is a finite A-module.

In the following proposition we shall prove that the singularities of a curve lying on a nonsingular *absolute surface* can be resolved by quadratic transformations applied to the surface. We shall base our demonstration on Zariski's proof of the algebro-geometric case; see Theorem 4 of [11].

PROPOSITION 5. *Let (R, M) be a two dimensional regular quotient ring of a point on an algebraic or absolute surface and let K be the quotient field of R. Let P be a minimal prime ideal in R and let w be the valuation of K with $R_w = R_P$. Let v be a valuation of K composite with w and having center M in R. Let (R_n, M_n) be the n-th quadratic transform of R along v and let $P_i = R_i \cap M_v$. Then P_i is a minimal prime ideal in R_i for $i = 1, 2, \cdots$; and there exists an integer n such that for any $i \geqq n$ we can choose a basis (x_i, y_i) of M_i for which $x_iR_i = P_i$, $v(y_i) = (0, 1)$, and $v(x) = (1, a)$ where a is some integer.*

Proof. Since $M_i \cap R = M$ and $P_i \cap R = R \cap M_w = P$ and since R_i is two dimensional, P_i must be a minimal prime ideal in R_i. Now let (x, y) be a basis of M, and suppose for instance that $v(x) \leqq v(y)$. Then $w(x) \leqq w(y)$. Therefore $x \notin P$; for otherwise we would have $w(x) > 0$ and hence $w(y) > 0$, i. e., $x \in R \cap M_w = P$ and $y \in R \cap M_w = P$, and hence $M = P$, which is a contradiction. Let $R^* = R[y/x]$, $M^* = R^* \cap M_v$, $P^* = R^* \cap M_w$. Let $\bar{R} = R/P$, $\bar{M} = M/P$, $\bar{R}^* = R^*/P^*$, $\bar{M}^* = M^*/P^*$, $\bar{R}_i = R_i/P_i$, $\bar{M}_i = M_i/P_i$, $\bar{K} = R_w/M_w$, $\bar{v} =$ the valuation of $\bar{K}$ induced by v, and $\bar{x}, \bar{y}$ the residue classes modulo M_w respectively of x, y. Then $(\bar{R}_i, \bar{M}_i)$ is a one dimensional local domain with quotient field $\bar{K}$, the real discrete valuation $\bar{v}$ of $\bar{K}$ has center $\bar{M}_i$ in $\bar{R}_i$, and $\bar{R} = \bar{R}_0 \subset \bar{R}_1 \subset \bar{R}_2 \subset \cdots$. Now $R_1 = R^*_{M^*}$, $M_1 = M_0R_1$ and $0 \neq \bar{x} \in \bar{R}$. Therefore $(\bar{x}, \bar{y})\bar{R} = \bar{M}$, $\bar{R}^* = \bar{R}[\bar{y}/\bar{x}]$, $\bar{M}^* = \bar{R}^* \cap M_{\bar{v}}$ and

$\bar{R}_1 = \bar{R}^*_{\bar{M}^*}$. Hence $\bar{M}\bar{R}^* = \bar{x}\bar{R}^*$ and $\bar{M}\bar{R}_1 = \bar{x}_1\bar{R}_1$. Similarly $\bar{M}_i\bar{R}_{i+1} = z_{i+1}\bar{R}_{i+1}$ with $z_{i+1} \varepsilon \bar{R}_{i+1}$ for $i = 0, 1, 2, \cdots$.

We shall now show that $\bigcup_{i=0}^{\infty} \bar{R}_i = R_{\bar{v}}$. Given c in $R_{\bar{v}}$ we can write $c = a/b$ with $0 \neq b$, $a \varepsilon \bar{R}$. If $b \not\varepsilon \bar{M}$ then $c \varepsilon \bar{R}$. Suppose that $b \varepsilon \bar{M}$. Since $\bar{v}(a) \geqq \bar{v}(b)$ we must have $a \varepsilon \bar{M}$. Hence $b = b_1 z_1$, $a = a_1 z_1$ with a_1, b_1 in $\bar{R}_1$. Since $c = a_1/b_1$, again either $c \varepsilon \bar{R}_1$ or $a_1, b_1 \varepsilon \bar{M}_1$. Suppose that $c \not\varepsilon \bar{R}_1$. Then $a_1 = a_2 z_2$, $b_1 = b_2 z_2$ with $a_2, b_2 \varepsilon \bar{R}_2$. Similarly if $c \not\varepsilon \bar{R}_{n-1}$ then $a = a_n z_1 z_2 \cdots z_n$ and $b = b_n z_1 z_2 \cdots z_n$ with $a_i, b_i \varepsilon \bar{R}_i$. Since $\bar{M}_{n-1} = \bar{R}_{n-1} \cap M_{\bar{v}}$, we must have $\bar{v}(z_n) > 0$, and hence that $\bar{v}(b) \geqq n$ where we take Z as the value group of the real discrete valuation $\bar{v}$. Since $\bar{v}$ is real discrete and since $b \neq 0$, for some n we must have $c \varepsilon \bar{R}_n$. Therefore $\bigcup_{i=0}^{\infty} \bar{R}_i = R_{\bar{v}}$. Let S_i be the integral closure of $\bar{R}_i$ in $\bar{K}$. Then S_i is a Dedekind domain with a finite number of prime ideals (§ 39 of [5]), and since the quotient ring of S_i with respect to any prime ideal is a real discrete valuation ring, we must have $S_i = \bigcap_{u \varepsilon W_i} R_u$ where W_i is a finite set of real discrete valuations of $\bar{K}$. It is clear that the valuations in W_i are exactly the valuations of $\bar{K}$ having center $\bar{M}_i$ in $\bar{R}_i$. Therefore $W_i \supset W_{i+1}$. Let $u_1, u_2, \cdots, u_h$ be the valuations in W_0 different from $\bar{v}$. Since $R_{\bar{v}}$ is a maximal subring of $\bar{K}$, we can find $a_i \varepsilon R_{\bar{v}}$ such that $a_i \not\varepsilon R_{v_i}$. Since $\bigcup_{j=0}^{\infty} \bar{R}_j = \bar{R}_{\bar{v}}$, $a_i \varepsilon \bar{R}_{m_i}$ for some integer m_i. Let $m = \max(m_1, m_2, \cdots, m_h)$. Then $a_i \not\varepsilon \bar{R}_m$ for $i = 1, 2, \cdots, h$. Therefore $W_m = \{\bar{v}\}$, i.e., $R_{\bar{v}}$ is the integral closure of $\bar{R}_m$. By Lemma 13, $R_{\bar{v}}$ is a finite module over $\bar{R}_m$. Hence by the Hilbert basis theorem, we can find n such that $\bar{R}_i = \bar{R}_{\bar{v}}$ whenever $i \geqq n$; from now on we shall assume that $i \geqq n$. Fix $\bar{y}_i$ in $\bar{R}_i$ with $\bar{v}(\bar{y}_i) = 1$. Let y_i be an element of R_i belonging to the residue class $\bar{y}_i$. Then $v(y_i) = (0, 1)$. By Lemma 9, P_i is a principal ideal; let $P_i = x_i R_i$. Since $(R_i)_{P_i} = R_w$, $v(x_i) = (1, a)$. Finally given z in M_i let $\bar{z}$ be the residue class of z modulo P_i. Then $\bar{z} \varepsilon \bar{M}_i$ and hence $\bar{z} = \bar{y}_i \bar{t}$ with $\bar{t} \varepsilon \bar{R}_i$. Fix t in R_i belonging to the residue class $\bar{t}$. Then $z - y_i t \varepsilon P_i$, i.e., $z \varepsilon (x_i, y_i)R_i$. Therefore $(x_i, y_i)R_i = M_i$.

Remark 3. Observe that Lemma 13, which is crucial in the proof of Proposition 5, breaks down for certain abstract local rings (regular of dimension two) as can be shown by using an example due to F. K. Schmidt; see page 24 of Zariski's paper [12]. To indicate this we shall use Zariski's notation. We have then two independent variables x and t over a certain field k of characteristic $p \neq 0$, $\Sigma' = k(x, t)$, o' is the valuation ring of a certain real discrete valuation v' of Σ'/k, $\Sigma = \Sigma'(\tau)$ with $\tau = t^{1/p}$ and o is the

one dimensional local domain $o'[\tau]$ with quotient field Σ. Let T be the integral closure of o' in Σ. Since o is integral over o', T is also the integral closure of o in Σ. Let $R = o'[X]$ and $P = (X^p - t)R$. Then it is easily verified that R is a two dimensional regular local domain, P is a minimal prime ideal in R, and $X \rightarrow \tau$ is an o'-homomorphism of R onto o with kernel P, i.e., o is isomorphic to R/P. Since T is not a finite o'-module (see part e on page 447 of [9]) and since o is a finite o'-module, T cannot be a finite o-module. Thus Lemma 13 is not applicable to the minimal prime ideal P of the two dimensional regular local domain R.

Now let K be the quotient field of R, let M be the maximal ideal in R, let w be the valuation of K with $R_w = R_P$, and let v be a valuation of K composed with w and having center M in R. Suppose, if possible, that Proposition 5 were true for the regular local domain (R, M). Then, in the notation of Proposition 5, we would have that R_n/P_n is a valuation ring. From the considerations made in the proof of Proposition 5, it follows that this would imply that o can be transformed into a regular local ring by applying to o, n successive quadratic transformations in the sense described on page 24 of [12], but this is impossible as is proved on pages 24-25 of [12]. Therefore, Proposition 5 is not applicable to the valuation v having center M in the regular two dimensional local domain R. Since Proposition 5 is essentially based on Lemma 13,[5] this again shows, as we have directly proved above, that Lemma 13 is not valid for arbitrary two dimensional regular local domains.

THEOREM 2.[6] Let (R, M) *be a two dimensional regular local domain with quotient field* K. *Let* v *be a valuation of* K *having center* M *in* R *and* R*-dimension zero. Let* f *be a given nonzero element of* R. *Then there exists a quadratic transform* (R^*, M^*) *of* R *along* v *and a basis* (x^*, y^*) *of* M^* *such that* $f = x^{*a}y^{*b}d$ *where* a *and* b *are nonnegative integers,* d *is a unit in* R^*, *and where the following conditions are satisfied:* (A) *If* v *is real of rational rank one, then* $b = 0$. (B) *If* v *is real of rational rank two then either* $b = 0$ *or* $v(x^*)$ *and* $v(y^*)$ *form an integral basis for the value group of* v. (C) *If* v *is of rank two and if* R *is the quotient ring of a point either on an algebraic surface or on an absolute surface, then* $v(x^*) = (1, h)$ *and* $v(y^*) = (0, 1)$ *where we are writing the* v*-values of elements of* K *as lexicographically ordered pairs of integers and where* h *is some integer.*

[5] I. e., the only fact about valuations centered in local domains which is used in the proof of Proposition 5 and in which the local domains are qualified to be either algebraic or absolute is Lemma 13.

[6] See Proposition 3 of [2].

Proof. First assume that v is real. Let $R_0 = R$ and $M_0 = M$. Let (x_0, y_0) be a basis of M_0 and let (R_i, M_i) be the i-th quadratic transform of R_0 along v. We shall define elements (x_i, y_i) of R_i by induction on i. Let then $i = m > 0$ and assume that we have defined x_i, y_i for $i = 1, 2, \cdots, m-1$. Suppose first that $v(y_{i-1}) \geqq v(x_{i-1})$. Let $S_i = R_{i-1}[y_{i-1}/x_{i-1}]$ and $N_i = M_{i-1}S_i$. Let $P_i = M_v \cap S_i$. Let z_i be the residue class modulo N_i containing y_{i-1}/x_{i-1}. Then by Lemma 3 of [1], $S_i/N_i = (R_{i-1}/M_{i-1})[z_i]$ and z_i is transcendental over R_{i-1}/M_{i-1}. Since P_i is a maximal ideal in S_i containing N_i, P_i/N_i must be a maximal ideal in S_i/N_i and hence $P_i/N_i = g_i(z_i)(S_i/N_i)$ where $g_i(X)$ is a monic irreducible polynomial in $(R_{i-1}/M_{i-1})[X]$. Let $G_i(X)$ be a monic polynomial in $R_{i-1}[X]$ which when reduced modulo M_{i-1} gives $g_i(X)$. We set $x_i = x_{i-1}$ and $y_i = G_i(y_{i-1}/x_{i-1})$. Secondly, if $v(y_{i-1}) < v(x_{i-1})$ then we set $x_i = y_{i-1}$ and $y_i = x_{i-1}/y_{i-1}$. Then by Corollary 1 of [1] $(x_i, y_i)R_i = M_i$.

Let $f_0 = f$ and define by induction $f_i \varepsilon R_i$ by the equation $f_{i-1} = x_i^{u_i} f_i$ where u_i is a nonnegative integer and where f_i is an element in R_i prime to x_i (by Lemma 9, R_i is a unique factorization domain). Let $f_{i,1}, f_{i,2}, \cdots, f_{i,h_i}$ be the irreducible factors of f_i in R_i and let $w_{i,j}$ be the valuation of K whose valuation ring is the quotient ring of R_i with respect to $f_{i,j}R_i$. Let W_i be the set of valuations u of K such that u has center M_i in R_i and u is composed with $w_{i,j}$ for some j. By Lemma 8, W_i is a finite set. For a given u in W_i, let $P_i = M_u \cap R_{i-1}$. Since u is nontrivial, $P_i \neq (0)$. Since $x_i \notin M_u \cap R_i$ and since $x_i R_i = M_{i-1}R_i$, $P_i \neq M_{i-1}$. Therefore P_i is a minimal prime ideal in R_{i-1}. Since R_{i-1} is a unique factorization domain and since $f_{i-1} \varepsilon P_i$, we must have $P_i = f_{i-1,j}R_{i-1}$ for some j, i.e., $u \varepsilon W_{i-1}$. Thus $W_i \subset W_{i-1}$. Since by Lemma 12, $\bigcup_{i=1}^{\infty} R_i = R_v$, since v is real, and since no element of W_i is real, it follows that $\bigcap_{i=1}^{\infty} W_i = \emptyset$. Since W_i is finite, $W_m = \emptyset$ for some m, i.e., f_m is a unit in R_m. Hence we have $f = x_m{}^A y_m{}^B D$ where A and B are nonnegative integer and D is a unit in R_m. If either v is of rational rank one or if v is of rational rank two and $v(x_m)$ and $v(y_m)$ are rationally dependent, then the proof can be completed by the argument of the last part of the proof of Proposition 3 of [2]. Now suppose that v is of rational rank two and that $v(x_m)$ and $v(y_m)$ are rationally independent. We may then take R^* to be R_m and $x^* = x_m$, $y^* = y_m$. It then remains to be shown that $v(x^*)$ and $v(y^*)$ form an integral basis for the value group of v. Let $v(x^*) = p$ and $v(y^*) = q$; and suppose for instance that $p < q$. Fix a representative system k in R^* of R^*/M^* (k is not in general a field; note that we take zero as the representative of M^*/M^*). Let z be an arbitrary nonzero element of R^* and let $v(z) = r$. Fix an integer n so that $r < np$. Then we can find

a polynomial $H(X,Y)=\sum_{i+j\leqq n} H_{ij}X^iY^j$ of degree at most n with coefficients H_{ij} in k such that $z^*=z-H(x^*,y^*)\,\varepsilon\, M^{*n}$ (see § 5 of [6]). Since $v(u)>r$ for any $u\,\varepsilon\, M^{*n}$, we must have that $v(z^*)>r$ and hence that $r=v(z)=v(H(x^*,y^*))$. Since p and q are rationally independent and since $v(H_{ij})=0$ whenever $H_{ij}\neq 0$, we can find $H_{st}\neq 0$ such that $v(H_{st}x^{*s}y^{*t})<v(H_{ij}x^{*i}y^{*j})$ whenever $H_{ij}\neq 0$ and whenever either $i\neq s$ or $j\neq t$. Therefore $r=v(z)=v(H(x^*,y^*))=v(H_{st}x^{*s}y^{*t})=sp+tq$. Therefore for any nonzero element z_1 of K, we must have $v(z_1)=s_1p+t_1q$ where s_1 and t_1 are integers. Thus we have shown that $\{p,q\}$ is an integral basis of the value group of v. This completes the proof of (A) and (B).

To prove (C), assume that v is of rank two, that R is the quotient ring of a point either on an algebraic or on an absolute surface and that we are writing the values of elements of K as lexicographically ordered pairs of integers. By Proposition 5, we can find a quadratic transform $(\bar{R},\bar{M})$ of R along v and a basis $\bar{x},\bar{y}$ of $\bar{M}$ such that $v(\bar{x})=(1,p)$ and $v(\bar{y})=(0,1)$ where p is some integer. Let $(\bar{R}_i,\bar{M}_i)$ be the i-th quadratic transform of $\bar{R}$ along v. Let $\bar{x}_i=\bar{x}/\bar{y}_i$ and $\bar{y}_i=\bar{y}$. Since $v(\bar{x})>iv(\bar{y})$ for any integer i, it follows that $(\bar{x}_i,\bar{y}_i)\bar{R}_i=\bar{M}_i$. Let $f=\bar{x}^ag_0$ where a is a nonnegative integer and where g_0 is an element of $\bar{R}$ prime to $\bar{x}$. Then $v(g_0)=(0,t)$ where t is some nonnegative integer. Define g_i by induction by the equation $g_{i-1}=\bar{y}_i^{e_i}g_i$ where e_i is a nonnegative integer and g_i is an element of $\bar{R}_i$ prime to $\bar{y}_i$. Since $\bar{M}_{i-1}\bar{R}_i=\bar{y}_i\bar{R}_i$, we have that $e_i>0$ whenever g_{i-1} is a nonunit in $\bar{R}_{i-1}$ (i.e., whenever $g_{i-1}\,\varepsilon\,\bar{M}_{i-1}$). Therefore, if g_{i-1} is a nonunit in $\bar{R}_{i-1}$ for $i=1,2,\cdots,n$, then $v(g_0)\geqq(0,n)$. Hence for some integer $m\leqq t$, g_m must be a unit in $\bar{R}_m$. Hence $f=x^{*a}y^{*b}d$ where b is a nonnegative integer, d is a unit in $R^*=\bar{R}_m$, $x^*=\bar{x}_m$, and $y^*=\bar{y}_m$.

We shall now prove a generalization to abstract two dimensional regular local domains of Zariski's theorem on the factorization of birational transformation between nonsingular algebraic surfaces into local quadratic transformations; see the lemma on page 538 of [11].

Theorem 3. *Let (R,M) and (R',M') be regular two dimensional local domains with a common quotient field K such that (R',M') has center M in R. Then R' is a quadratic transform of R.*

We shall precede the proof proper of this theorem by three preparatory lemmas.

Lemma 14. *Let A be a normal domain with quotient field K such that A contains a unique maximal ideal P and let $0\neq x\,\varepsilon\, K$ such that $x\notin A$ and*

$(1/x) \notin A$. *Let* $B = A[x]$ *and let* $M = PB$. *Then:* (1) *there exists a valuation* v *of* K *such that* $R_v \supset B$, $M_v \cap B = M$; (2) $M \cap A = P$; (3) *the residue class* $\bar{x}$ *modulo* M *containing* x *is transcendental over* A/P, *i.e.,* R_v/M_v *contains the polynomial ring* $(R/M)[\bar{x}]$ *and hence* R_v/M_v *is of positive transcendence degree over* R/M.

Proof. Follows by the considerations made by Zariski on pages 26-27 of [14].

LEMMA 15. *Let* (R, M) *be an* n *dimensional regular local domain with quotient field* K *and* $n > 1$. *Let* $\{x_1, x_2, \cdots, x_n\}$ *be a minimal basis of* M. *Let* $A = R[x_2/x_1, x_3/x_1, \cdots, x_n/x_1]$ *and* $P = MA$. *Then:* (1) $P = x_1 A$; (2) P *is a minimal prime ideal of* A; (3) A_P *is the valuation ring of a real discrete valuation* w *of* K; (4) $A/P = (R/M)[\bar{y}_2, \bar{y}_3, \cdots, \bar{y}_n]$ *where* $\bar{y}_i =$ *the residue class modulo* P *containing* x_i/x_1, *and* $\bar{y}_2, \bar{y}_3, \cdots, \bar{y}_n$ *are algebraically independent over* R/M; (5) $R_w/M_w = (R/M)(\bar{y}_2, \bar{y}_3, \cdots, \bar{y}_n)$; *and* (6) w *is the "M-adic divisor of R," i.e.,* w *is completely defined by setting* $w(a/b)$ $=$ *(leading degree of* a*)* $-$ *(leading degree of* b*) where* a *and* b *are any two nonzero elements of* R, *and* w *is of* R*-dimension* $(n-1)$.

Proof. (1) is obvious, (2) and (4) are proved in Lemma 3 of [1]. Now let N be a maximal ideal in A containing P, let $R^* = A_N$ and $M^* = NR^*$. Then by Corollary 1 of Lemma 3 of [1], R^* is a regular local domain and $P^* = y_1 R^*$ is a minimal prime ideal of R^*. Hence $R^*_{P^*} = R_w$ where w is a real discrete valuation of K. Now (3) follows since $A_P = R^*_{P^*}$, and (5) follows from (4) and the fact that $A_P/(PA_P)$ is the quotient field of A/P. Now let a be an arbitrary nonzero element of R and let d be the leading degree of a. Then $a = f(x_1, x_2, \cdots, x_n)$ where $f(X_1, X_2, \cdots, X_n)$ is a form of degree d with coefficients in R not all in M; and hence $a = x_1{}^d g$ where $g = f(1, x_2/x_1, x_3/x_1, \cdots, x_n/x_1) \notin P$ since $\bar{y}_2, \bar{y}_3, \cdots, \bar{y}_n$ are algebraically independent over R/M; i.e., $w(a) = d$. Finally, it follows by (5) that w is of R-dimension $(n-1)$ and this proves (6).

LEMMA 16. *Let* (R, M) *and* (R', M') *be as in Theorem* 3 *and assume that* $R \neq R'$. *Then there exists a nonunit* z *in* R' *such that* $MR' \subset zR'$.

Proof. Suppose, if possible, that MR' is primary for M'. Let $(\bar{R}, \bar{M})$ and $(\bar{R}', \bar{M}')$ be the completions of R and R' respectively. Let w be the M'-adic divisor of R'. Then w has center M in R and R-dimension of $w \geqq R'$-dimension of $w = 1$. Therefore by Theorem 1, w is a prime divisor also for R. Hence by Theorem 2 of [13], (R, M) and (R', M') are both subspaces

of (R_w, M_w) and hence R is a subspace of R'. Therefore we may canonically assume that $\bar{R}$ is a subring and a subspace of $\bar{R}'$. By Theorem 1, R_w/M_w is finitely generated over R/M and hence R'/M' must be a finite algebraic extension of R/M. Since $\bar{R}'/\bar{M}' = R'/M'$ and $\bar{R}/\bar{M} = R/M$, we have that $\bar{R}/\bar{M}$ is a finite algebraic extension of R/M. Also MR' is primary for M' implies that $\bar{M}'\bar{R}'$ is primary for $\bar{M}'$. Therefore by Theorem 8 of [3], $\bar{R}'$ is integral over $\bar{R}$. Let E and E' be the quotient fields of $\bar{R}$ and $\bar{R}'$ respectively. Let A be the set of valuations v of K having center M in R and let B be the set of valuations u of K for which $R_u \supset R$. Let v be a member of A. Then by Lemma 13 of [2], v has an extension $\bar{v}$ to E having center $\bar{M}$ in $\bar{R}$. Since $\bar{R}'$ is integral over $\bar{R}$, we must have $R_{\bar{v}'} \supset \bar{R}'$ for some extension $\bar{v}'$ of $\bar{v}$ to E'. Since $R_v = R_{\bar{v}'} \cap K$ and since $R' = \bar{R}' \cap K$ (Lemma 2 of [2]), we must have $R' \subset R_v$. Therefore $R' \subset \bigcap_{v \in A} R_v$. Now let u be a member of B for which $P = M_u \cap R \neq M$. Let $\bar{K} = R_u/M_u$, $\bar{R} = R/P$ and $\bar{M} = M/P$. Let $\bar{u}$ be a valuation of $\bar{K}$ having center $\bar{M}$ in $\bar{R}$ and let u^* be the valuation of K which is composed of u and $\bar{u}$. Then $R_{u^*} \supset R_u$ and u^* has center M in R. Therefore $R' \subset \bigcap_{u \in B} R_u = R$, i.e., $R' = R$. Thus our assumption that MR' is primary for M' is absurd. Since R' is two dimensional, $MR' \subset p$ where p is a minimal prime ideal in R' and by Lemma 9, $p = zR'$ with $z \in R'$; i.e., $MR' \subset zR'$.

Proof of Theorem 3. If $R = R'$ then there is nothing to prove. So assume that $R < R'$. By Lemma 9, there exists a discrete rank two valuation v of K having center M' in R'. By Theorem 1, R-dimension of $v = 0 = R'$-dimension of v. Let $\{x, y\}$ be a minimal basis of M, and suppose for instance that $v(x) \leqq v(y)$. Let $t = y/x$. If $1/t \in R'$ then $t \in R_v$ and $1/t \in R_v$ and hence $1/t \notin M_v$, i.e., $1/t \notin M'$ and hence $t \in R'$. Now suppose, if possible, that $t \notin R'$. Then $1/t \notin R'$. Let $A' = R'[t]$ and $P' = M'A'$. By Lemma 14, there exists a valuation w of K with center P in A, and

$$R_w/M_w \supset A'/P' = (R'/M')[\bar{t}]$$

where $\bar{t}$ is the residue class modulo P' containing t and $\bar{t}$ is transcendental over R'/M'. Let $A = R[t]$, $P = MA$ and $P^* = P' \cap A$. Then $P^* \cap R = M$ and hence t is *a fortiori* transcendental over R modulo M. Hence $P^* = MA = P = xA$. Therefore $A_P \supset R_w$ and hence by Lemma 15, w must be the M-adic divisor of R. Since $w(x) = 1$ and since $w(M') > 0$, we have that $x \in M'$ and $x \notin (M')^2$, i.e., that x is an irreducible nonunit in R'. By Lemma 16, $x = az$ and $y = bz$ where a and b are in R' and z is a nonunit in R'. Since by Lemma 9, R' is a unique factorization domain, we must have $y = cx$

with c in R', i.e., $t = c \varepsilon R'$. Thus our assumption that $t \notin R'$ is absurd. Let (R_i, M_i) be the i-th quadratic transform of R along v. Then $R_1 \subset R'$ and R' has center M_1 in R_1. Similarly, if $R_1 < R'$ then $R_2 \subset R'$ and R' has center M_2 in R_2 and so on. Suppose, if possible, that $R_i \subset R'$ for all i. Then by Lemma 12, $R_v \subset R'$ and hence $R_v = R'$. Since v is rank two, R_v, i.e., R' is nonnoetherian which is absurd. Therefore, for some integer n, we must have $R_n = R'$.

LEMMA 17. *Lemma* 12 *remains true if we replace the assumption that R_{i+1} is a quadratic transform of R_i for $i = 0, 1, 2, \cdots$, by the weaker assumption that $M_{i+1} \cap R_i = M_i$ for $i = 0, 1, 2, \cdots$.*

Proof. Follows by Lemma 12 and Theorem 3.

Appendix.*

That an ordered abelian group G of finite rank can be embeded (as an ordered subgroup) in the lexicographically ordered direct sum $R^{(\rho)}$ of ρ copies of the additive group R of real numbers can be proved as follows: First we prove that *if G is rationally complete* (i.e., with any element g of G and an arbitrary nonzero integer n, G always contains an element g^* such that $ng^* = g$) *then we have the stronger result that G itself can be expressed as a lexicographically ordered direct sum of ordered subgroups of rank one,* i.e., if $0 = G_0 < G_1 < \cdots < G_\rho = G$ is the sequence of isolated subgroups of G then G contains ordered subgroups $H_1, H_2, \cdots, H_\rho$ of rank one such that $G_{\rho-i} = H_{i+1} \oplus H_{i+2} \oplus \cdots \oplus H_\rho$ where the sum is lexicographically ordered direct (i.e., every element g of $G_{\rho-i}$ has a unique expression $g = h_{i+1} + h_{i+2} + \cdots + h_\rho$ with h_j in H_j, if $g^* = h^*_{i+1} + h^*_{i+2} + \cdots + h^*_\rho$ with h^*_j in H_j is any other element of $G_{\rho-i}$ different from g and if $h_{i+j} = h^*_{i+j}$ for $j = 1, 2, \cdots, t-1$ and $h_{i+t} \neq h^*_{i+t}$ then $g < g^*$ or $g > g^*$ according as $h_{i+t} < h^*_{i+t}$ or $h_{i+t} > h^*_{i+t}$). For $\rho = 1$ this is obvious and to apply induction, assume that $\rho > 1$ and that the assertion is true for $\rho - 1$. Let H be the maximal isolated subgroup of G. Then H is of rank $\rho - 1$, G/H is of rank one, and H and G/H are both rationally complete. By the induction hypothesis, $H = H_2 \oplus H_3 \oplus \cdots \oplus H_\rho$ where the sum is lexicographically ordered direct. Let B be a rationally independent rational basis of G/H and for each b in B fix β in G contained in the residue class b. Let H_1 be the set of all elements of G which depend rationally on the β's. Then it can be easily verified that

* Received February 2, 1956.

H_1 is order isomorphic to G/H and that $G = H_1 \oplus H = H_1 \oplus H_2 \oplus \cdots \oplus H_\rho$ where the direct sums are lexicographically ordered.

In the general case when G is not necessarily rationally complete, it is enough to observe that G can always be embeded in a rational completion i. e. in an abelian group G^* which is rationally complete and for any element g^* of G we have $ng^* \varepsilon G$ for some nonzero integer n. The ordering of G can be uniquely extended to G^* and then G^* is again of rank ρ. As shown above, we can find in G^* subgroups $H_1, H_2, \cdots, H_\rho$ of rank one such that $G^* = H_1 \oplus H_2 \oplus \cdots \oplus H_\rho$ where the direct sum is lexicographically ordered. Now each H_i being archimedian is order isomorphic to a subgroup of R and hence G^* is order isomorphic to a subgroup of $R^{(\rho)}$. Therefore G is order isomorphic to a subgroup of $R^{(\rho)}$. Now we shall show that if G is not rationally complete then it need not be order isomorphic to a lexicographically ordered direct sum of rank one groups. To show this, let G denote the torsionfree abelian group studied by L. Pontrijagin in Example 2 of Appendix 1 of his paper: The theory of topological commutative groups; *Annals of Mathematics*, Vol. 35 (1934), pp. 361-388. Then G is of rational rank two and it is not expressible as a direct sum of rational rank one subgroups. Let (u, v) be a rational basis of G. Then each element g of G has a unique expression $g = au + bv$ where a and b are rational numbers. Let $g^* = a^*u + b^*v$ be any other element of G where a^* and b^* are rational numbers. We set $g > g^*$ either if $a > a^*$ or if $a = a^*$ and $b > b^*$. This turns G into an ordered abelian group and we have $\rho(G) = 2$. Suppose if possible that G is the lexicographically ordered direct sum of two ordered rank one subgroups H_1 and H_2. Then we have that $r(H_1) = r(H_2) = 1$ and hence that G is the direct (group theoretic) sum of two rational rank one subgroups; this is absurd.

We observe that if in the above proof we replace simple induction by transfinite induction then we obtain the following more general result: Let G be an ordered abelian group and let $S(G)$ be the family of isolated subgroups of G simply ordered by inclusion. Assume that $S(G)$ is well ordered. Let G^* be a rational completion of G and extend the ordering of G to G^* (uniquely). Then G^* is order isomorphic to a lexicographically ordered direct sum of a well ordered family T [which is order isomorphic to $S(G)$] of archimedian groups.

HARVARD UNIVERSITY.

REFERENCES.

[1] S. Abhyankar and O. Zariski, "Splitting of valuations in extensions of local domains," *Proceedings of the National Academy of Sciences*, vol. 41 (1955), pp. 84-90.

[2] ———, "Local uniformization on algebraic surfaces over ground fields of characteristic $p \neq 0$," forthcoming in the *Annals of Mathematics*.

[3] I. S. Cohen, "On the structure and ideal theory of complete local rings," *Transactions of the American Mathematical Society*, vol. 59 (1946), pp. 54-106.

[4] W. Krull, "Allgemeine Bemertungstheorie," *Jour. für Mathematik*, vol. 167 (1932), pp. 160, 196.

[5] ———, *Idealtheorie*, Berlin (1935).

[6] ———, "Dimensionstheorie in Stellenringen," *Jour. für Mathematik*, vol. 179 (1938), pp. 204-226.

[7] ———, "Zur Theorie der kommutativen Integritätsbereiche," *ibid.*, vol. 192 (1954), pp. 230-252.

[8] S. MacLane and O. F. G. Schilling, "Zero-dimensional branches of rank one on algebraic varieties," *Annals of Mathematics*, vol. 40 (1939), pp. 507-520.

[9] F. K. Schmidt, "Über die Erhaltung der Kettensätze der Idealtheorie bei beliebigen endlichen Körpererweiterungen," *Mathematische Zeitschrift*, vol. 41 (1936), pp. 443-440.

[10] O. Zariski, "Foundations of a general theory of birational correspondences," *Transactions of the American Mathematical Society*, vol. 53 (1943), pp. 490-542.

[11] ———, "Reduction of the singularities of algebraic three dimensional varieties," *Annals of Mathematics*, vol. 45 (1944), pp. 472-542.

[12] ———, "The concept of a simple point of an abstract algebraic variety," *Transactions of the American Mathematical Society*, vol. 62 (1947), pp. 1-52.

[13] ———, "A simple analytical proof of a fundamental property of birational transformations," *Proceedings of the National Academy of Sciences*, vol. 35 (1949), pp. 62-66.

[14] ———, "Applicazioni geometriche della teoria delle valutazioni," *Rendiconti del Circolo Matematico di Palermo*, Ser. 5, vol. 13 (1954), pp. 1-38.

ON FRENET'S EQUATIONS.*

By AUREL WINTNER.

1. If Γ is an oriented, rectifiable Jordan arc in the X-space, where $X=(x,y,z)$, let $\Gamma \varepsilon C''$ mean that the vector function $X(s)$ has a continuous second derivative with respect to the arc length s. Thus $U_1=X'$, where $'=d/ds$, is a continuously differentiable unit vector. Let κ denote the (non-negative, continuous) function $|U_1'|$ of s, the curvature on Γ.

Let Γ be called a *Frenet curve* (in symbols: $\Gamma \varepsilon F$) if it is a $\Gamma \varepsilon C''$ corresponding to which there exists a vector function U_3 possessing the following properties: U_3 is a unit vector which, at every s, is orthogonal to U_1 and has a continuous derivative U_3' which is linearly dependent on the vector product $[U_3, U_1]$.

2. The following considerations will deal with the class $\Gamma \varepsilon F$. The emphasis will be two-fold: on the one hand, $\Gamma \varepsilon F$ does not involve the existence of a third derivative for the vector function $X(s)$, where Γ: $X=X(s)$ and, on the other hand, $\Gamma \varepsilon F$ allows the vanishing of the curvature $\kappa(s)=|X''(s)|$ (on arbitrarily complicated s-sets, which must, of course, be closed sets, since $\kappa(s)$ is continuous). It is therefore unexpected that $\Gamma \varepsilon F$ proves to be the natural condition for the existence of a Frenet theory. In fact, it turns out that

(I) *every* $\Gamma \varepsilon F$ *possesses a unique, continuous torsion* $\tau=\tau(s)$

and that

(II) *Frenet's equations are valid on every* $\Gamma \varepsilon F$.

(I) and (II) imply the existence and the continuous differentiability of a principal normal and of a binormal on every $\Gamma \varepsilon F$. *But these unit vectors in Frenet's equations need not be unique* (although, by (I), the torsion *is* unique), not even if the differentiability assumption $\Gamma \varepsilon C''$, contained in the assumption $\Gamma \varepsilon F$, is refined to the existence of arbitrarily high derivatives of $X(s)$; simply because Γ can contain segments of straight lines (in a finite or infinite number).

The generality under which (I) and (II) are secured by it (Sections 3-4

* Received July 27, 1955.

below) is not, however, the only merit of the *curve* class $\Gamma \varepsilon F$. In fact, its principal merit consists in its ability to deal with certain desiderata in the theory of *surfaces*. This will be explained and carried out in Sections 5 and 6-7, respectively.

3. Let a Γ satisfy the conditions required of a $\Gamma \varepsilon F$ in Section 1. Define $U_2 = U_2(s)$ by placing $U_2 = [U_3, U_1]$. Then (U_1, U_2, U_3) is a continuously differentiable matrix function of s, and is an orthogonal matrix, of deterimnant $+1$.

By the last of the requirements which define the condition $\Gamma \varepsilon F$, the (continuous) vector U_3' is a scalar multiple of U_2. Let this scalar factor be denoted by $-\tau$, and let $\tau = \tau(s)$ be declared to be the torsion of Γ. Then, since $U_3' = -\tau U_2$ is continuous, as is the unit vector U_2, it is clear that τ is continuous. It also follows from $|U_2| = 1 \neq 0$ that τ is uniquely determined by U_3' and U_2 and, therefore, by U_2. But $U_2 = U_2(s)$ is not uniquely determined by Γ when $\kappa(s) = 0$. The following consideration will, however, show that τ is unique at every s, as claimed by (I), Section 2 (in fact, $U_3' = -\tau U_2$ implies that $\tau = -U_3 \cdot U_2'$, whereas U_2 and/or U_3 becomes indeterminate only at points contained in the *interior* of such straight line segments as may be contained in Γ; but $\tau = \tau(s)$ proves to be $\equiv 0$ on any such segment).

Let (F_i), where $i = 1, 2, 3$, denote the i-th of the relations

$$U_1' = \kappa U_2, \qquad U_2' = -\kappa U_1 + \tau U_3, \qquad U_3' = -\tau U_2.$$

Then (F_3) holds by the definition of τ. If (F) denotes the system of the three relations (F_i), then the assertion of (II), Section 2, is that (F) holds at every point s of every $\Gamma \varepsilon F$. This assertion will be proved in Section 4.

4. Suppose first only that $\Gamma \varepsilon C''$. Then there exists a continuous $\kappa = \kappa(s) \geqq 0$, defined as the length of the continuous vector function X''. Denote by $B_1, B_2, \cdots$ the (finite or infinite) sequence of open s-intervals on which $U_1'(s) \neq 0$, where $U_1 = X'$. The s-interval, say A, which is the topological map of the entire arc Γ is supposed to be closed. It can be assumed that $A - B$, where $B = B_1 + B_2 + \cdots$, is not vacuous. For otherwise $\kappa \equiv 0$ on A, hence Γ is a line segment, and so all three equations (F_i) become satisfied by placing $\tau \equiv 0$, $U_3 = [U_1, U_2]$ and choosing U_2 to be any constant unit vector which is orthogonal to $U_1 = \text{const}$. Correspondingly, it will be clear that it is sufficient to prove (II), Section 2, under the assumption that Γ contains no line segments.

Then the interval A is the closure of the open set B. But $U_2 = U_1'/\kappa$ defines a $U_2 = U_2(s)$ on every component B_i of B. Since κ and U_1' are

defined and continuous on A, and since B is dense on A, it is clear that it is possible to define on A a unique continuous U_2 satisfying (F_1), the first of the three equations (F).

On the other hand, (F_3) is true by definition (Section 3). Finally, (F_2) follows by inserting (F_1) and (F_3) into the derivative of $U_2 = [U_3, U_1]$.

5. Let $\Gamma \varepsilon F^*$ mean that Γ belongs to the subclass F^* of the class F which is restricted by the hypothesis that $\kappa(s) \neq 0$ on Γ. The books of analytically-minded authors do not fail to emphasize that the (continuous) torsion, which they define as $1/\kappa^2$ times the determinant of the first three derivatives of X, is left undefined unless $\kappa(s) \neq 0$ on Γ *and* $X(s)$ possesses a (continuous) third derivative. The resulting geometrical interpretation of the torsion (in terms of the binormal spherical image) is classical, of course. The analytical advantages of this geometrical approach were noticed, however, only recently, in [2], pp. 770-774.

It was shown in [4], pp. 243-244, that a $\Gamma \varepsilon C''$ satisfies the C^3-assumption of the traditional treatment if and only if it is a $\Gamma \varepsilon F^*$ on which $\kappa(s) > 0$, instead of being just continuous, is continuously differentiable, a hypothesis which prevents the formulation of important geometrical situations, for which $\Gamma \varepsilon F^*$ turns out to be the natural assumption (cf. [4], p. 246, pp. 247-249 and p. 257). This holds not only for a curve Γ as such but also for classical curves Γ drawn on the surface, such as asymptotic lines and geodesics (cf. [2], pp. 773-774, and [3], pp. 608-610).

It is clear, however, that not only the traditional presentation (C^3) but also its geometrical refinement (F^*) excludes every straight line Γ, the (otherwise) isolated zeros s of $\kappa(s)$ if $\Gamma \varepsilon A$, and closed s-sets (which can be Cantor sets of positive measure) if $S \varepsilon C^\infty$, where A and C^∞ denote the classes of curves Γ for which the function $X(s)$ is analytic and possesses derivatives of arbitrarily higher order, respectively. This leads to the anomalous situation that, for instance, such fundamental facts as center around the concept of geodesic torsion or the Beltrami-Enneper theorem have never been formulated for those points of Γ at which $\kappa(s) = 0$.

This was the reason for replacing the F^*-class by the F-class, as introduced in Section 1. It remains to be shown that the class $\Gamma \varepsilon F$ is not too inclusive to be useful in disposing of the anomalies referred to before.

6. Let n be a positive integer and S a (sufficiently small, open, simply connected piece of a) surface in the X-space, where $X = (x, y, z)$. Then an $S \varepsilon C^n$ is defined by the property that S has some parametrization S: $X = X(u, v)$ in which all partial derivatives of the function $X(u, v)$ which have a (collective) order not exceeding n exist, are continuous, and those of

the first order are such that their vector product $[X_u, X_v]$ does not vanish at any point of the (u,v)-domain under consideration.

The most immediate use of the concept of a continuous torsion, as defined in Section 3, is that the following assertion becomes true without any restriction:

(i) *If* Γ *is a geodesic on an* $S \varepsilon C^2$, *then* $\Gamma \varepsilon F$.

In order to prove (i), let $N = N(s)$ denote the (oriented) unit normal vector of S along Γ. Since $S \varepsilon C^2$ and since Γ is a geodesic, $\Gamma \varepsilon C''$. Hence $U_1 = X'$ has a continuous derivative. The same is true of U_2 and U_3, if the latter two vectors are defined to be N and $[U_1, U_2] = [X', N]$, respectively. But the definition of a geodesic Γ requires the vanishing of $[X', N]' \cdot X'$ along Γ, whilst $[X', N]' \cdot [X', N]$ vanishes identically (since $[X', N] \cdot [X', N] = 1$). This means that U_3' is orthogonal to U_1 as well as to U_3 and is, therefore, a scalar multiple of U_2. Hence, (i) follows from the definition of the class $\Gamma \varepsilon F$ (Section 1).

In answering a question I raised some time ago, Professor Hartman ([1], pp. 724-726) has shown that if D is a direction through P within T, where P denotes a point of any $S \varepsilon C^2$ and T the plane tangent to S at P, then (P, D) determines a *unique* geodesic of S. Hence, (i) supplies the following result:

(ii) *If* $S \varepsilon C^2$, *then every point of* S *and a direction through it* (*on* S) *define a unique geodesic torsion,* the latter being defined as the torsion (at P), supplied by (i), of the geodesic determined by (P, D).

If $\Gamma \varepsilon C'$ means that Γ is an oriented, rectifiable Jordan arc for which $X(s)$ possesses a continuous first derivative, then a corollary of (ii) can be formulated as follows:

(iii) *A* $\Gamma \varepsilon C'$ *on an* $S \varepsilon C^2$ *is a line of curvature of* S *if and only if the geodesic torsion of* Γ *vanishes identically* (*provided that* S *is free of umbilical points*).

In fact, since (ii) assures the existence of a geodesic torsion along Γ, it is only necessary to repeat the considerations of [3], pp. 608-609, in order to obtain (iii).

7. Let $S \varepsilon C^3$. Then the Gaussian curvature K and the coefficients of the second fundamental form $Ldu^2 + 2Mdudv + Ndv^2$ (when referred to a C^3-parametrization $X = X(u,v)$ of S) are continuously differentiable functions of (u,v). Suppose that K is negative on S. Then $LN - M^2 < 0$. Hence the identical vanishing of the second fundamental form along a curve

Γ of S represents for Γ either of two systems of ordinary differential equations and, if either of them is written in the form $du/dt = f(u,v)$, $dv/dt = g(u,v)$, then both functions f, g are continuously differentiable and do not contain the independent variable, t. Hence the solutions $u = u(t)$, $v = v(t)$ have continuous second derivatives with respect to t. Since an asymptotic line Γ of S is defined by $\Gamma: X = X(u(t), v(t))$, it follows that $\Gamma \varepsilon C''$. It will be shown that $\Gamma \varepsilon C''$ can be improved to $\Gamma \varepsilon F$:

(iv) *If $K < 0$ on an $S \varepsilon C^3$, then there exists a continuous torsion on every asymptotic line Γ of S; in fact, $\Gamma \varepsilon F$. In addition, the square of the torsion is identical with $-K$ on S.*

First, since $S \varepsilon C^3$ implies that $S \varepsilon C^2$, it follows from the fact $\Gamma \varepsilon C''$ (which could *not* have been concluded from just $S \varepsilon C^2$) that the unit normal $N = N(s)$ to S has a continuous first derivative along Γ, as does $U_1 = X'(s)$, where Γ: $X = X(s)$. Hence, if U_2 and U_3 are defined to be $[N, X']$ and N, respectively, then (U_1, U_2, U_3) is a continuously differentiable orthogonal matrix (of determinant $+1$). But the definition of an asymptotic line Γ, used above, requires the vanishing of $X' \cdot N'$ along Γ, whilst $N \cdot N'$ vanishes identically (since $N \cdot N = 1$). This means that $U_3' = N'$ is orthogonal to U_1 as well as to U_3 and is, therefore, a scalar multiple of U_2. Hence the assertion $\Gamma \varepsilon F$ of (iv) follows from the definition of the Frenet class F (Section 1).

The remaining assertion of (iv), that concerning the value of $-K$ along Γ (Beltrami-Enneper), can be concluded from the Frenet system (F), belonging to $U_3 = N$, in the usual way. Cf. the corresponding remarks in [2], p. 773, which deal with (iv) under the (by now superfluous) hypothesis that the curvature $\kappa(s)$ does not vanish on the asymptotic line Γ. Actually, the last assertion of (iv) then follows from the first assertion of (iv) without the assumption $\kappa(s) > 0$ also, simply for reasons of continuity (in fact, if the trivial case $\kappa(s) \equiv 0$ is disregarded, then the zeros $s = s_0$ of $\kappa(s)$ are cluster points of an open s-set on which $\kappa(s) > 0$). Cf. also (III) in [3], p. 609.

It may finally be mentioned that, in view of [5], pp. 858-859, the assumption, $S \varepsilon C^3$, in (iv) appears to be necessary and sufficient in order that an $S \varepsilon C^2$ of negative K be such as to possess a C^2-parametrization S: $X = X(u, v)$ in which the parameter lines, $u =$ const., $v =$ Const. are asymptotic lines of S.

8. If $\Gamma \varepsilon C'''$ means that Γ: $X = X(s)$ is a rectifiable Jordan arc for which the vector function $X(s)$ possesses a continuous third derivative, then the following criterion (*) holds:

(*) *A* $\Gamma \varepsilon F$ *is a* $\Gamma \varepsilon C'''$ *if and only if its curvature* $\kappa(s)$ *is continuously differentiable.*

This is a generalization of a criterion in [4], pp. 243-244, where (*) was proved under the hypothesis $\kappa(s) \neq 0$ (that is, under the assumption that $\Gamma \varepsilon F$ is strengthened to $\Gamma \varepsilon F^*$ in the sense of Section 5). It is, however, clear from the proof (*loc. cit.*) and from the definition of the Frenet class F (Section 1) that the first of the assertions of (*), the assertion in which the continuous differentiability of $\kappa(s)$ is the assumption, holds also if $\Gamma \varepsilon F^*$ is relaxed to $\Gamma \varepsilon F$.

In order to prove the second assertion of (*), suppose that $\Gamma \varepsilon C'''$ and let A and $B = B_1 + B_2 + \cdots$ be defined as in Section 4. Then $U_1 = X'$ has a continuous second, hence $\kappa^2 = X''.X''$ a continuous first, derivative on A. Actually, $\kappa = \kappa(s) \geqq 0$ itself must have a continuous first derivative on A. In order to prove this at an arbitrary point s_0 of A, two cases must be distinguished, according as s_0 is in B or in $A - B$. In the first case, the assertion is trivial, since $\kappa(s_0) > 0$. Let therefore $\kappa(s_0) = 0$. Then, if (F_1) is applied at s_0 and at a nearby s, it is seen that $\kappa(s) U_2(s)$ is identical with the difference $U_1'(s) - U_1'(s_0)$. But the ratio of the latter to $s - s_0$ tends, as $s \rightarrow s_0$, to a limit, since $U_1(s)$ has a second derivative. Consequently, $\kappa(s) U_2(s)/(s - s_0)$ must tend, as $s \rightarrow s_0$, to a limit. Since $U_2(s) \rightarrow U_2(s_0) \neq 0$, this means that $\kappa(s)/(s - s_0)$ has a limit. In view of $\kappa(s_0) = 0$, this proves the differentiability of $\kappa(s)$ at s_0, and therefore at every point of A. Finally, the continuity of the derivative $\kappa'(s)$ follows from the circumstance that, since $U_1''(s)$ is continuous, the preceding limit process holds uniformly.

It is clear from Section 2 that (*) can be interpreted as a criterion supplying, in terms of the behavior of the curvature, a necessary and sufficient condition in order that a $\Gamma \varepsilon C''$ be a $\Gamma \varepsilon C'''$, provided that Γ has a continuous torsion. In fact, this proviso is contained in (and, when combined with $\Gamma \varepsilon C''$, becomes equivalent to) the hypothesis $\Gamma \varepsilon F$ of (*).

Appendix.

With reference to a Γ: $X(s)$ of class F, let R denote either sheet $(0 < t < \infty$ or $-\infty < t < 0)$ of the ruled surface generated by the binormal of F; so that R: $X = X(t, s) \equiv t U_3(s)$, where $t \neq 0$ and $X = (x, y, z)$. Thus $X(t, s)$ is a function of class C^1 and

$$[X_t, X_s] = [U_3, t U_3'] = - t\tau [U_3, U_2],$$

by (F_3). Hence the unit normal vector, say $M = M(t, s)$, of R exists if $\tau \neq 0$. In fact, $\pm M$ is seen to be $[U_2, U_3] = U_1$. Accordingly, if $\Gamma \varepsilon F$ and $\tau \neq 0$, then R is a surface of class C^1. It turns out, however, that *R is a surface*

of class C^2 (even though its "geometrical" parametrization, R: $X = tU_3(s)$, cannot in general be of class C^2, since $U_3(s) \varepsilon C^2$ is surely not true if only $\Gamma \varepsilon F$, or for that matter $\Gamma \varepsilon C^3$, is assumed). This can be seen as follows:

The surface $R \varepsilon C^1$ has a C^1-parametrization in terms of which the unit normal M is a function of class C^1. Hence, $R \varepsilon C^2$ can be concluded by an elementary argument (based on the classical theorem on local implicit functions) which was repeatedly applied in the writings of Hartman and myself (cf., e. g., [*3*], pp. 368-369, where, incidentally, the issue involved, that of the preservation of the C^2-character of a surface $S \varepsilon C^2$ under a "parallel" deformation, is very similar to the present construction of the surface R from the curve Γ, where $\Gamma \varepsilon F$, hence $\Gamma \varepsilon C''$). Of course, $R \varepsilon C^2$ means that *some* parametrization $X = X(u;v)$ of R is of class C^2 (with $[X_u, X_v] \neq 0$). The point is that, against expectation, the "geometrical" parametrization, that in which $(u,v) = (t,s)$, fails to be such a parametrization in general.

Needless to say, the ruled surface R is a torse, in the sense of having a normal $M = M(t,s)$ which is independent ($= \pm U_1$) of the position t on any generating line ($s =$ const.) of R. Since $R \varepsilon C^2$, this implies* that the Gaussian curvature K of R vanishes identically; cf. the end of the footnote below. The result is therefore as follows:

(a) *If* $\tau \neq 0$ *on a* $\Gamma \varepsilon F$ (*where* $\kappa \neq 0$ *is not assumed*), *then the ruled surface* $R = R(\Gamma)$: $X = tU_3(s)$, *where* $0 < |t| < \infty$, *is a torse* ($K \equiv 0$) *of class* C^2.

---

* Certain difficulties inherent to Euler's definition of a torse are known since Lebesgue's thesis ([*2*], pp. 319-342). But it may not be necessary to go to such extremes as Lebesgue went (continuous but not one-to-one parametrizations) in order to show that the theory of torses is not as simple as it appears from the texts of differential geometry, including the rigor-conscious books. For is it true that if $K \equiv 0$ on an $S \varepsilon C^2$, then a neighborhood of every point of S can be "ruled," so as to be a torse in Euler's sense also? I can neither prove nor believe this, not even under the assumption $S \varepsilon C^\infty$ which, in view of the possibility of clustering zeros of H (i. e., of "flat" points, where $H^2 = 0 = K$), is hardly stronger than $S \varepsilon C^2$ (in view of Theorem (†), p. 134, of [*1*], there is no trouble when a torse $S \varepsilon C^2$ is free of flat points, but there could be trouble even if there is just one such point). A counterexample, with $S \varepsilon C^\infty$, would be the first such instance in the differential geometry of surfaces as to require the full force of (function-theoretical) analyticity (or at least quasi-analyticity, rather than just C^∞-character).

The converse inference, that in which the identical vanishing of the Gaussian curvature is the assertion, can be concluded without any calculation (whenever the surface is of class C^2). In fact, if S is a torse in Euler's synthetic sense of the term, then the normal image of each of the generating lines is a single point. Since the C^2-character of S suffices to justify the applicability of Fubini's theorem on product measures, it thus becomes clear that the normal image of any subset of S is of measure zero on the unit sphere.

This result, (α), has a counterpart, (β), for the case in which the binormal U_3 is replaced by the tangent U_1 of Γ (but not for the case of the remaining U_i, the principal normal; the hindrance being that equation (F_2), in contrast to (F_3) and (F_1), does not contain just one U on the right). The counterpart is as follows:

(β) *If $\kappa \neq 0$ on a $\Gamma \varepsilon F$ (where $\tau \neq 0$ is not assumed), i.e., if $\Gamma \varepsilon F^*$, then the ruled surface $P = P(\Gamma): tU_1(s)$, where $0 < |t| < \infty$, is a torse ($K \equiv 0$) of class C^2.*

In fact, what corresponds to the last formula line when R is replaced by P is

$$[X_t, X_s] = [U_1, tU_1'] = t\kappa[U_1, U_2],$$

by (F_1). Hence it is clear that (β) follows by a repetition of the proof of (α).

It will be noted that the proof of (α) or (β) succeeds because the ruled surfaces traditionally attached to a Γ: $X = X(s)$, which are the ruled surfaces $X(s) + tU_i(s)$, are reduced to $tU_i(s)$, i.e., that the base curve, instead of being Γ, is made to be 0.

THE JOHNS HOPKINS UNIVERSITY.

REFERENCES.

[1] P. Hartman, "On the local uniqueness of geodesics," *American Journal of Mathematics*, vol. 72 (1950), pp. 723-730.

[2] ——— and A. Wintner, "On the fundamental equations of differential geometry," *ibid.*, vol. 72 (1950), pp. 757-774.

[3] ——— and A. Wintner, "On geodesic torsions and parabolic and asymptotic curves," *ibid.*, vol. 74 (1952), pp. 607-625.

[4] A. Wintner, "On the infinitesimal geometry of curves," *ibid.*, vol. 75 (1953), pp. 241-259.

[5] ———, "On indefinite binary Riemannian metrics," *ibid.*, vol. 77 (1955), pp. 853-867.

APPENDIX.

[*1*] P. Hartman and A. Wintner, "On the curvatures of a surface," *American Journal of Mathematics*, vol. 75 (1953), pp. 127-141.

[*2*] H. Lebesgue, "Integrale, longueur, aire," *Annali di Matematica Pura ed Applicata*, ser. 3, vol. 7 (1902), pp. 231-359.

[*3*] A. Wintner, "On parallel surfaces," *American Journal of Mathematics*, vol. 74 (1952), pp. 365-376.

THE STRUCTURE OF FACTORS OF AUTOMORPHY.*

By R. C. GUNNING.

One of the principal tools in the study of automorphic functions of one and several complex variables is the representation of an automorphic function as the quotient of two holomorphic functions whose zeros are invariant under the group of transformations acting, although the functions themselves are not invariant; these auxiliary functions, that is, satisfy a relation of the form $f(Tz) = \nu_T(z)f(z)$, where $\nu_T(z)$ are holomorphic and nowhere vanishing functions called factors of automorphy. Such auxiliary functions were first studied in connection with elliptic functions in one variable and abelian functions in several variables, where they have been called the Jacobi or intermediary functions. Their importance in deriving the Riemann conditions on the period matrix of a multi-torus, in discussing the divisors on a multi-torus, and in proving the existence theorems of abelian functions, has long been recognized. Appell was the first to prove, in several variables, that the particular factors of automorphy involved in the Jacobi functions are, in a sense, the most general possible factors on the multi-torus [1]. Corresponding auxiliary functions have been used in other cases also, in connection with the Poincaré and Eisenstein series, but, as Bochner has pointed out [3], no systematic attempts have been made to give a general classification of these factors corresponding to the work of Appell. The present paper contains the proofs of some results announced earlier [11] in connection with this classification problem, and a discussion of the significance of this classification as a generalization of that of the Jacobi functions.

The method utilized in the proofs is the application of potential theory in the form of the theory of harmonic differential forms on Kaehler manifolds; recent investigations of this subject, in connection with analytic manifolds and transcendental algebraic geometry, have yielded a wealth of applicable results. Although some knowledge of the terminology and notation used in the study of complex manifolds is presupposed, the relevant topological and differential-geometric results are collected in the first two sections. It has been possible to develop the subject on this foundation without requiring any

* Received December 3, 1955.

extensive algebraic or topological background; the presentation given here uses this approach for the benefit of those whose interest lies in the function-theoretic aspects of the subject. A supplementary section, Section 5, illustrates an algebraic development of the classification theorem.

I would like to express my warmest thanks here to Professor Bochner for suggesting this problem to me, and for many valuable discussions on this and related topics.

I. Introductory.

1. Suppose that D is a simply-connected complex analytic manifold of complex dimension n, and that Γ is a countable group of analytic homeomorphisms of D onto itself subject to the following restrictions:

(1.1) Each point $z \in D$ lies in a coordinate neighborhood U such that either $TU = U$ or $TU \cap U = \emptyset$ for any $T \in \Gamma$.

(1.2) For each pair of points $z_1, z_2 \in D$ either $\Gamma z_1 = \Gamma z_2$ or there exist coordinate neighborhoods U_1, U_2 of z_1, z_2 respectively such that $\Gamma U_1 \cap \Gamma U_2 = \emptyset$. (Here, $\Gamma z = \bigcup_{T \in \Gamma} Tz$.)

(1.3) The quotient space D/Γ, with the quotient topology, is a compact space.

(1.4) There is a Kaehler metric on D which is invariant under the action of the group Γ.

For any subset $W \subset D$ let Γ_W be the subgroup of Γ consisting of all transformations T for which $TW = W$; thus whenever W is a coordinate neighborhood, Γ_W is a properly discontinuous group of analytic homeomorphisms on a bounded affine subdomain. Applying the standard methods of the theory of automorphic functions on bounded affine domains to the particular coordinate neighborhoods (1.1), one sees that the subgroups Γ_z are finite for each z, and that there are arbitrarily small coordinate neighborhoods of each point which satisfy (1.1); in particular there is a coordinate neighborhood U of each point $z \in D$ such that for any $T \in \Gamma$, either $TU \cap U = \emptyset$ or $TU = U$ and $Tz = z$.

The collection of all points of D left fixed by a transformation $T \in \Gamma$ other than the identity forms a proper analytic subvariety S_T of D. The set

$$S = \bigcup_{T \in \Gamma - I} S_T,$$

which one notes is a Γ-invariant analytic subvariety of D, is called the *singular set* of D; its image S/Γ under the canonical projection $\rho : D \to D/\Gamma$

is called the *singular set* of D/Γ. The mapping $\rho: D - S \to D/\Gamma - S/\Gamma$ is then a local homeomorphism, defining the structure of a complex analytic manifold on $D/\Gamma - S/\Gamma$ and exhibiting $D - S$ as a regular covering manifold. Select a fixed base point $z_0 \in D - S$, and for each $T \in \Gamma$ select a path α_T from z_0 to Tz_0 in $D - S$; the uniquely defined homotopy class of the loop $\rho(\alpha_T)$ in $D/\Gamma - S/\Gamma$ will be denoted by $\tilde{\psi}(T)$, so that $\tilde{\psi}$ is the standard isomorphism of Γ onto the fundamental group $\pi_1(D/\Gamma - S/\Gamma)$ of $D/\Gamma - S\Gamma$ based at z_0. The injection of $D/\Gamma - S/\Gamma$ into D/Γ induces a further homomorphism of $\pi_1(D/\Gamma - S/\Gamma)$ onto $\pi_1(D/\Gamma)$, and the composite of this homomorphism with $\tilde{\psi}$ defines a homomorphism ψ of Γ onto $\pi_1(D/\Gamma)$. One can see that the kernel of this homomorphism is the normal subgroup Γ_0 of Γ generated by all transformations which possess fixed points, so that $\Gamma/\Gamma_0 \cong \pi_1(D/\Gamma)$; furthermore, Γ_0 is contained in the normal subgroup of Γ generated by all transformations which are of finite order. Any homomorphism of Γ into a commutative field therefore vanishes on Γ_0 and on all commutators of elements of Γ, so determines a homomorphism on $H_1(D/\Gamma)$; for the field of real numbers in particular, the so-called universal coefficient theorem [7] asserts that each such homomorphism corresponds to a one-dimensional cohomology class on D/Γ with real coefficients.

Suppose that the group Γ is presented in some fixed manner as the factor group of a finitely generated free group H modulo a group of relations P, which is also finitely generated. The free generators of H will be denoted by $t_1, \cdots, t_p$, and their images under the canoncial projection homomorphism $\mathrm{H} \to \Gamma$ will be denoted by $T_1, \cdots, T_p$ respectively. The subgroup $[\mathrm{P}, \mathrm{H}]$ of H generated by all words of the form $s^{-1}r^{-1}sr$ for arbitrary $r \in \mathrm{P}$, $s \in \mathrm{H}$, is called the *commutator* of P in H; replacing P by H itself, one defines correspondingly the *commutator subgroup* $[\mathrm{H}, \mathrm{H}]$ of H. Note that $[\mathrm{P}, \mathrm{H}]$ is a normal subgroup of H, hence also of the group $\mathrm{P} \cap [\mathrm{H}, \mathrm{H}]$ defined as the set-theoretic intersection of P and $[\mathrm{H}, \mathrm{H}]$. Letting $H_2(D)$, $H_2(D/\Gamma)$ denote the singular homology groups of D and D/Γ, the projection $\rho: D \to D/\Gamma$ induces a homomorphism $\rho_*: H_2(D) \to H_2(D/\Gamma)$ whose image will be denoted by $S_2(D/\Gamma)$. We shall next construct explicitly a useful homomorphism

$$\Phi: \mathrm{P} \cap [\mathrm{H}, \mathrm{H}]/[\mathrm{P}, \mathrm{H}] \to H_2(D/\Gamma)/S_2(D/\Gamma), \tag{2}$$

related to a theorem of H. Hopf [13].

An element $v \in \mathrm{P} \cap [\mathrm{H}, \mathrm{H}]$ is uniquely expressible as a reduced word in the free generators of H, say

$$v = t_{\alpha(q)}{}^{\epsilon(q)} \cdots t_{\alpha(1)}{}^{\epsilon(1)},$$

where $\epsilon(\nu)=\pm 1$. Let $j(1)$ be the smallest integer for which

$$t_{\alpha(j(1))}{}^{\epsilon(j(1))}=t_{\alpha(1)}{}^{-\epsilon(1)};$$

let $j(2)$ be the smallest integer distinct from 1 and $j(1)$ such that

$$t_{\alpha(j(2))}{}^{\epsilon(j(2))}=t_{\alpha(2)}{}^{-\epsilon(2)}; \text{ etc.}$$

Thus there is associated to the word v a unique one-to-one mapping j of a subset of the integers $1,\cdots,q$ onto the complementary set. For each r in the domain of the mapping j construct an arbitrary singular 1-simplex $\sigma_r{}^1$ in $D-S$ such that

$$\partial\sigma_r{}^1=T_{\alpha(j(r))}{}^{\epsilon(j(r))}\cdots T_{\alpha(1)}{}^{\epsilon(1)}z_0-T_{\alpha(r-1)}{}^{\epsilon(r-1)}\cdots T_{\alpha(1)}{}^{\epsilon(1)}z_0.$$

Introduce also the singular 1-simplex $\bar{\sigma}_r{}^1=-T_{\alpha(r)}{}^{\epsilon(r)}\sigma_r{}^1$ and the singular 1-chain $\kappa^1(v)=\sum_r(\sigma_r{}^1+\bar{\sigma}_r{}^1)$, where the summation is extended over all r in the domain of the mapping j. It is clear from the construction that $\partial\kappa^1(v)=0$ and that $\rho\kappa^1(v)=0$, where $\rho:D\to D/\Gamma$ is again the canonical projection. Since D is simply-connected, there is a singular 2-chain $\kappa^2(v)$ such that $\partial\kappa^2(v)=\kappa^1(v)$, and hence $\partial\rho\kappa^2(v)=\rho\partial\kappa^2(v)=0$. Let $\phi(v)$ be the coset of the homology class of $\rho\kappa^2(v)$ in $H_2(D/\Gamma)/S_2(D/\Gamma)$. Two types of arbitrary choices were involved in the definition of the elements $\phi(v)$, namely the selections of the $\sigma_r{}^1$ and of the $\kappa^2(v)$; clearly the result is independent of these choices, so that ϕ is a well-defined mapping. One sees immediately that it is even a homomorphism of $\mathrm{P}\cap[\mathrm{H},\mathrm{H}]$ into $H_2(D/\Gamma)/S_2(D/\Gamma)$.

One can further see that the homomorphism ϕ is actually independent of the choice of the auxiliary mapping j, so long as j is one-to-one between a subset of the indexing set and its complement, and $t_{\alpha(r)}{}^{\epsilon(r)}=t_{\alpha(j(r))}{}^{-\epsilon(j(r))}$. As a first corollary of this independence, one notes that $[\mathrm{P},\mathrm{H}]$ lies in the kernel of the homomorphism ϕ; the induced homomorphism Φ on the factor group $\mathrm{P}\cap[\mathrm{H},\mathrm{H}]/[\mathrm{P},\mathrm{H}]$ is the desired homomorphism.

In case Γ has no fixed points, D/Γ is itself a complex analytic manifold and D is its universal covering manifold. The mapping Φ is then an isomorphism onto; indeed, as a second corollary of the independence of Φ from the auxiliary mapping j, one sees that Φ coincides with the isomorphism established by Hopf. The group $S_2(D/\Gamma)$ may also be interpreted in this case as the image of the second homotopy group of D/Γ in the second homology group under the canonical mapping.

2. Some differential-geometric properties of the spaces D and D/Γ are also needed as background for the subsequent discussion. These are all

known theorems for manifolds without singularities; the fact that they remain true for manifolds with singularities of the types admitted here in the spaces D/Γ when Γ contains transformations with fixed points has been verified in detail by W. Baily [2]. For the notation to be used in the sequel, as well as for definitions of the fundamental concepts involved, see for example [10, 18].

By *de Rham's theorems,* the cohomology groups of the space D/Γ with complex coefficients are isomorphic to the d-cohomology groups of Γ-invariant complex differential forms on the manifold D, where d is the exterior differentiation operator on D. *Hodge's theorem* asserts the existence of a unique Δ-harmonic Γ-invariant differential form representing each cohomology class. Since the manifold D is complex-analytic, the ordinary exterior differentiation may be decomposed as the sum of two complex operators, $d = \partial + \bar{\partial}$. One may define a harmonic operator $\square$, associated to the differentiation $\bar{\partial}$, analogous to the Laplace-Beltrami operator Δ; since D is Kaehler, one even has $\square = \frac{1}{2}\Delta$. Analogs of the Hodge theorem may be obtained for the operator $\square$. A crucial step in this development is the *decomposition theorem:* every Γ-invariant C^∞ differential form ϕ of type (p,q) on D which satisfies $\bar{\partial}\phi = 0$ may be written in the form $\phi = \bar{\partial}\psi + \theta$, where ψ is a Γ-invariant differential form of type $(p, q-1)$ and θ is a Γ-invariant $\square$-harmonic form of type (p,q), in the sense that $\square\theta = 0$. From the Kaehler hypothesis, one secures $d\theta = 0$. From this follows a topological invariance theorem: if b^r is the r-th Betti number of the space D/Γ and $h^{p,q}$ is the dimension of the complex linear space of Γ-invariant $\square$-harmonic differential forms of type (p,q), then $b^r = \sum_{p+q=r} h^{p,q}$.

The Γ-invariant $\square$-harmonic differential forms $\omega_\alpha(z)$ of type $(1,0)$ on D are called the *abelian differentials* (of the first kind) on D/Γ; these may be defined equivalently as the Γ-invariant differential forms of type $(1,0)$ which satisfy either $\bar{\partial}\omega_\alpha(z) = 0$ or $d\omega_\alpha(z) = 0$. Since the forms $\omega_\alpha(z)$ are closed, it is possible to introduce in addition the *abelian integrals* $w_\alpha(z) = \int^z \omega_\alpha(\zeta)$; these may be characterized as the set of all single-valued holomorphic functions on D satisfying $w_\alpha(Tz) = w_\alpha(z) + \hat{\omega}_\alpha(T)$, for all $T \in \Gamma$, where $\hat{\omega}_\alpha(T)$ are complex constants called the *periods* of the integral. If $\{\omega_\alpha(z)\}$ form a basis for the complex linear space of abelian differentials on D/Γ, then since the operator $\square$ is real on a Kaehler manifold, the conjugate differentials $\{\overline{\omega}_\alpha(z)\}$ form a basis for the complex linear space of $\square$-harmonic differential forms of type $(0,1)$; applying the topological invariance theorem, in the case $r = 1$, there are precisely $\frac{1}{2}b^1$ linearly independent abelian differentials on D/Γ. Let $\{T_j\}$ be a set of transformations of Γ representing the infinite cyclic generators of the abelianized group $\Gamma/[\Gamma,\Gamma]$; from our previous geo-

metrical considerations, there are at most b^1 such elements. These may be called a *basis* for the group Γ. Setting $\omega_{\alpha j} = \hat{\omega}_\alpha(T_j)$, the matrix $\Omega = (\omega_{\alpha j})$ is called the *period matrix* corresponding to the bases $\{\omega_\alpha(z)\}$ and $\{T_j\}$. Since the mapping $\hat{\omega}_\alpha : T \to \hat{\omega}_\alpha(T)$ is a homomorphism of Γ into the additive group of complex numbers, it vanishes on elements of finite order and on the subgroup $[\Gamma, \Gamma]$; hence all the periods $\hat{\omega}_\alpha(T)$ of the abelian differential $\omega_\alpha(z)$ are integer linear combinations of the basic periods listed in row α of the period matrix. Since there can exist no abelian differential with purely imaginary periods, one notes that there are precisely b^1 basic transformations, and that the b^1 by b^1 *matrix* $\begin{pmatrix} \Omega \\ \bar{\Omega} \end{pmatrix}$ is non-singular.

II. Classification of Factors of Automorphy.

3. The collection of all holomorphic functions on the complex analytic manifold D form an abelian group $\mathfrak{A}(D)$ under addition. The group Γ then has a natural interpretation as a group of operators on $\mathfrak{A}(D)$, where the action of an element $T \in \Gamma$ on $f(z) \in \mathfrak{A}(D)$ is defined to yield $f(Tz)$.

Definition. A *summand of automorphy* for the group Γ on D is a mapping σ of Γ into $\mathfrak{A}(D)$, the image of an element $T \in \Gamma$ being denoted by $\sigma_T(z)$, such that $\sigma_{ST}(z) = \sigma_S(Tz) + \sigma_T(z)$. In algebraic terminology, an equivalent restatement of this definition is that a summand of automorphy is a one-cocycle of Γ with coefficient group $\mathfrak{A}(D)$ [8].

LEMMA 1. *There exists for any summand of automorphy σ a C^∞ function $f(z)$ on D such that*

$$f(Tz) = f(z) + \sigma_T(z). \tag{3}$$

Proof. Select a finite number of pairs of open coordinate neighborhoods $V_j \subset \bar{V}_j \subset U_j$ of D such that the sets $\rho(V_j)$ cover D/Γ, that for any $T \in \Gamma$ either $TU_j = U_j$ or $TU_j \cap U_j = \phi$, and that the subgroups $\Gamma_j = \Gamma_{U_j}$ are of orders $m_j < \infty$. For each j construct a real C^∞ function $\tilde{\mu}_j(z)$ on D such that $0 \leqq \tilde{\mu}_j(z) \leqq 1$, that $\tilde{\mu}_j(z) = 0$ for $z \notin U_j$, and that $\tilde{\mu}_j(z) = 1$ for $z \in V_j$. Then the functions

$$\mu_j(z) = \sum_{T \in \Gamma} \tilde{\mu}_j(Tz) / \sum_{T \in \Gamma} \sum_j \tilde{\mu}_j(Tz)$$

are Γ-invariant C^∞ functions on D, $\mu_j(z) = 0$ whenever $z \notin \bigcup_{T \in \Gamma} TU_j$, and $\sum_j \mu_j(z) \equiv 1$.

The functions $\sigma_{jT}(z) = \mu_j(z) \sigma_T(z)$ form a set of C^∞, although not holomorphic, summands of automorphy. Let

$$f_j(z) = \begin{cases} -1/m_j \sum_{S \in \Gamma_j} \sigma_{jS}(z) & \text{for } z \in U_j, \\ f_j(T^{-1}z) + \sigma_{jT}(T^{-1}z) & \text{for } z \in TU_j, \\ 0 & \text{otherwise.} \end{cases}$$

Then $f_j(z)$ is clearly C^∞ on D and satisfies $f_j(Tz) = f_j(z) + \sigma_{jT}(z)$. The function $f(z) = \sum_j f_j(z)$ is then the function whose existence was to be demonstrated.

In general there will not exist a holomorphic function satisfying (3). Select any C^∞ function $f(z)$ such that $f(Tz) = f(z) + \sigma_T(z)$. Since $\sigma_T(z)$ are holomorphic, $\bar\partial\sigma_T(z) = 0$ and $\bar\partial f(Tz) = \bar\partial f(z)$; thus $\bar\partial f(z)$ is a Γ-invariant $\bar\partial$-closed C^∞ differential form of type $(0,1)$. Applying the decomposition theorem of Section 2, $\partial f(z) = \bar\partial f^*(z) + \theta_\sigma(z)$, where $f^*(z)$ is a Γ-invariant C^∞ function on D, and $\theta_\sigma(z)$ is a Γ-invariant harmonic differential form of type $(0,1)$ on D. The form $\theta_\sigma(z)$, or the one-dimensional cohomology class $\hat\theta_\sigma$ it represents, is an obstruction associated to the summand $\sigma_T(z)$. It depends only upon the summand of automorphy $\sigma_T(z)$. For if $f_1(z)$ is another C^∞ function satisfying (3), and $\bar\partial f_1(z) = \bar\partial f_1^*(z) + \theta_\sigma^1(z)$, set $f_0(z) = f(z) - f^*(z) - f_1(z) + f_1^*(z)$ and $\theta_\sigma^0(z) = \theta_\sigma(z) - \theta_\sigma^1(z)$; then $f_0(Tz) = f_0(z)$ and $\bar\partial f_0(z) = \theta_\sigma^0(z)$ is harmonic, which implies $\theta_\sigma^0(z) = 0$. Clearly there will exist a holomorphic function satisfying (3) if and only if the obstruction $\theta_\sigma(z)$ of the summand $\sigma_T(z)$ vanishes; this in turn can be achieved by a simple modification of the original summand.

THEOREM 1. *If σ is any summand of automorphy for the group Γ on D, there is a unique homomorphism $\hat a: T \to a_T$ of Γ into the additive group of real numbers for which there will exist a holomorphic function $g(z)$ satisfying $g(Tz) = g(z) + \sigma_T(z) + 2\pi i\, a_T$.*

Proof. Replacing the function $f(z)$ considered in the above paragraph by $f(z) - f^*(z)$ if necessary, one may assume that $f(Tz) = f(z) + \sigma_T(z)$ and that $\theta_\sigma(z) = \bar\partial f(z)$ is a harmonic differential form. Thus

$$0 = d\theta_\sigma(z) = \partial\bar\partial f(z) = -\bar\partial\partial f(z) = -d\partial f(z),$$

so that $\partial f(z)$ is a closed, holomorphic differential form of type $(1,0)$. Since D is simply-connected, $g(z) = \int^z \partial f(\zeta)$ is a well-defined holomorphic function on D for which $dg(z) = \partial f(z)$. Consequently

$$g(Tz) = g(z) + \sigma_T(z) + 2\pi i\, b_T \tag{4}$$

for some complex constants b_T.

The mapping $T \to b_T$ is clearly a homomorphism of Γ into the additive group of complex numbers, so that $b_T = 0$ whenever T lies in the commutator subgroup of Γ, or is an element of finite order. As a result, this homomorphism may be considered as defining a one-dimensional cohomology class $\hat{b}_\sigma$ of D/Γ with complex coefficients. One already has the one-dimensional cohomology class $\hat{\theta}_\sigma$ defined by the secondary obstruction of the summand σ. For any 1-cycle α of the space D/Γ, represented by a singular 1-chain of D with boundary $Tz_0 - z_0$, one secures

$$\begin{aligned}\hat{\theta}_\sigma(\alpha) &= \int_{z_0}^{Tz_0} \theta_\sigma(z) = \int_{z_0}^{Tz_0} \bar{\partial} f(z)\\ &= \int_{z_0}^{Tz_0} df(z) - \partial f(z)\\ &= f(Tz_0) - g(Tz_0) - f(z_0) + g(z_0)\\ &= -2\pi i\, b_{T_0} = -2\pi i\, \hat{b}_\sigma(\alpha);\end{aligned}$$

therefore $\hat{b}_\sigma = \frac{1}{2\pi i}\hat{\theta}_\sigma$.

The most general holomorphic function satisfying an equation of the desired type (4), with perhaps different complex constant terms a_T appearing, is given by

$$g^*(z) = g(z) + \sum_{\alpha=1}^{\frac{1}{2}b^1} \xi_\alpha w_\alpha(z) + a$$

where a and ξ_α are complex constants and $\{w_\alpha(z)\}$ form a basis for the abelian integrals on D/Γ. Letting $\{T_j\}$ be a basis for the group Γ in the sense of Section 2 and $\Omega = (\omega_{\alpha j})$ be the corresponding period matrix, all of the constants a_T are expressible as linear combinations, with integer coefficients, of the numbers

$$a_{T_j} = b_{T_j} + \sum_{\alpha=1}^{\frac{1}{2}b^1} \xi_\alpha \omega_{\alpha j}.$$

Since the matrix $\begin{pmatrix}\Omega\\ \bar{\Omega}\end{pmatrix}$ is non-singular, values ξ_α may be selected so that a_{T_j}, and hence all a_T, are real numbers. Then the function $g^*(z)$ is obviously unique up to the additive constant a, while the homomorphism $T \to a_T$ is completely unique. This homomorphism defines a one-dimensional cohomology class $\hat{a}_\sigma$ of D/Γ with real coefficients. Letting $\hat{\omega}_\alpha$ be the one-dimensional cohomology classes of D/Γ defined by the abelian differentials $\omega_\alpha(z)$,

$$\hat{a}_\sigma = -\frac{1}{2\pi i}\hat{\theta}_\sigma + \sum_{\alpha=1}^{\frac{1}{2}b^1} \xi_\alpha \hat{\omega}_\alpha.$$

The remarks made during the course of the preceding proof have given a significant further interpretation to the homomorphism $T \rightarrow a_T$ by considering the one-dimensional cohomology class $\hat{a}_\sigma$ defined by this homomorphism, by way of the isomorphism $\Gamma/\Gamma_0 \cong \pi_1(D/\Gamma)$ established in Section 1. The cohomoloy classes $\hat{\omega}_\alpha$ form a subgroup of $H^1(D/\Gamma, C)$ which may be called the subgroup of *analytic cohomology classes.* The class $\hat{a}_\sigma$ may then be described as the unique real cohomology class in that coset of the subgroup of analytic cohomology classes which contains the obstruction $-\frac{1}{2\pi i}\hat{\theta}_\sigma$ of the summand σ. Also of interest is the cohomology class $\chi_2(\sigma) = \exp 2\pi i\, \hat{a}_\sigma$ with coefficients in the one-dimensional unitary group.

4. The collection of all holomorphic, nowhere-vanishing functions on the complex manifold D form an abelian group $\mathfrak{M}(D)$ under multiplication. The group Γ then has a natural interpretation as a group of operators on $\mathfrak{M}(D)$, where the action of an element $T \in \Gamma$ on $f(z) \in \mathfrak{M}(D)$ is defined to yield $f(Tz)$.

Definition. A *factor of automorphy* for the group Γ on D is a mapping ν of Γ into $\mathfrak{M}(D)$, the image of an element $T \in \Gamma$ being denoted by $\nu_T(z)$, such that $\nu_{ST}(z) = \nu_S(Tz)\nu_T(z)$. The set of all factors of automorphy form an abelian group under multiplication.

In algebraic terminology, a factor of automorphy is a one-cocycle of the group Γ with coefficients in the group $\mathfrak{M}(D)$. Two such cocycles or factors μ and ν are *cohomologous* if there exists a holomorphic, nowhere-vanishing function $h(z)$ on D such that $h(Tz)/h(z) = \mu_T(z)/\nu_T(z)$. We shall develop a classification of these cohomology classes analogous to that given in Theorem 1 for the corresponding additive cocycles, or summands of automorphy. It is more convenient, and actually more natural, to approach this problem in two steps, the first of which consists in the study of a weaker classification of factors of automorphy. For this purpose, one recalls that a *character* of the group is a homomorphism $\hat{c}: T \rightarrow c_T$ of Γ into the multiplicative group of complex numbers of modulus 1, the one-dimensional unitary group. Two factors of automorphy μ and ν are *equivalent* if there exists a character $\hat{c} = \{c_T\}$ of the group Γ and a holomorphic non-vanishing function $h(z)$ on D such that $\mu_T(z) = c_T\nu_T(z)\, h(Tz)/h(z)$.

For any group H and any abelian group G, let $\mathrm{Hom}(\mathrm{H}; G)$ be the group of all homomorphisms of H into G; for any subgroups $\mathrm{H}_1 \subset \mathrm{H}$ and $G_1 \subset G$, let $\mathrm{Hom}(\mathrm{H}_1, \mathrm{H}; G_1, G)$ be the subgroup of $\mathrm{Hom}(\mathrm{H}_1; G_1)$ consisting of all elements which can be extended to homomorphisms of H into G. As in

Section 1, the group Γ is assumed presented in some manner as the quotient of a finitely generated free group H modulo a group of relations P; the free generators of H are $t_1, \cdots, t_p$, and their images in Γ are $T_1, \cdots, T_p$ respectively.

If ν is any factor of automorphy, select a branch of the logarithm $\sigma_{t_j}(z) = \log \nu_{T_j}(z)$ for each free generator t_j. After this selection has been made, a unique holomorphic function $\sigma_t(z)$ can be associated to each $t \in \mathrm{H}$ by requiring that $\sigma_{st}(z) = \sigma_s(Tz) + \sigma_t(z)$. For every $t \in \mathrm{H}$, $\nu_T(z) = \exp \sigma_t(z)$; consequently whenever $t \in \mathrm{P}$

$$\hat{\sigma}(t) = \frac{1}{2\pi i}\sigma_t(z) \tag{5}$$

is an integer. Furthermore for $s, t \in \mathrm{P}$, $\hat{\sigma}(st) = \hat{\sigma}(s) - \hat{\sigma}(t)$, while for $t \in \mathrm{P}$ and $v \in \mathrm{H}$, $\hat{\sigma}(vtv^{-1}) = \frac{1}{2\pi i}[\sigma_v(TV^{-1}z) + \sigma_t(V^{-1}z) - \sigma_v(V^{-1}z)] = \hat{\sigma}(t)$; this therefore defines an element $\hat{\sigma} \in \mathrm{Hom}(\mathrm{P}/[\mathrm{P},\mathrm{H}]; Z)$, where Z is the additive group of integers. If $\sigma^*_{t_j} = \log \nu_{T_j}(z)$ are defined by selecting different branches of the logarithms, then for any $t \in \mathrm{H}$, $\sigma^*_t(z) - \sigma_t(z) = 2\pi i\, m_t$ for some integer m_t; the mapping $\hat{m}: t \to m_t$ is a homomorphism of H into the additive group Z. That is, for any two homomorphisms $\hat{\sigma}$ and $\hat{\sigma}^*$ determined by the same factor of automorphy,

$$\hat{\sigma}^* - \hat{\sigma} \in \mathrm{Hom}(\mathrm{P}/[\mathrm{P},\mathrm{H}], \mathrm{H}/[\mathrm{P},\mathrm{H}]; Z, Z). \tag{6}$$

LEMMA 2. *If* H *is a finitely generated free group and* P *is a normal subgroup, then there is a canonical isomorphism into*

$$\phi: \frac{\mathrm{Hom}(\mathrm{P}/[\mathrm{P},\mathrm{H}]; Z)}{\mathrm{Hom}(\mathrm{P}/[\mathrm{P},\mathrm{H}], \mathrm{H}/[\mathrm{P},\mathrm{H}]; Z, \bar{Q})} \to \mathrm{Hom}(\mathrm{P} \cap [\mathrm{H},\mathrm{H}]/[\mathrm{P},\mathrm{H}]; Z),$$

where Z is the additive group of integers and $\bar{Q}$ is a group containing the additive group of rational numbers.

Proof. An element $\sigma \in \mathrm{Hom}(\mathrm{P}/[\mathrm{P},\mathrm{H}]; Z)$ may be considered as a homomorphism of P into Z such that $\sigma([\mathrm{P},\mathrm{H}]) = 0$. Let $\phi(\sigma)$ be the restriction of σ to the subgroup $\mathrm{P} \cap [\mathrm{H},\mathrm{H}] \subset \mathrm{P}$; this defines a canonical homomorphism ϕ of $\mathrm{Hom}(\mathrm{P}/[\mathrm{P},\mathrm{H}]; Z)$ into $\mathrm{Hom}(\mathrm{P} \cap [\mathrm{H},\mathrm{H}]/[\mathrm{P},\mathrm{H}]; Z)$. Obviously $\mathrm{Hom}(\mathrm{P}/[\mathrm{P},\mathrm{H}], \mathrm{H}/[\mathrm{P},\mathrm{H}]; Z, \bar{Q})$ lies in the kernel of ϕ. Conversely suppose that $\phi(\sigma) = 0$, or what is the same, that $\sigma(\mathrm{P} \cap [\mathrm{H},\mathrm{H}]) = 0$; to complete the proof it is only necessary to show that σ can be extended to a homomorphism of H into Q. Let P be generated by $u_1, u_2, \cdots$, and write

$$u_j = t_1^{m_{j1}} \cdots t_p^{m_{jp}} s_j$$

for some $s_j \in [\mathrm{H}, \mathrm{H}]$ and integers m_{jk}. Values $\sigma(u_j)$ are given, and values $x_k = \sigma(t_k)$ are to be selected in such a manner that σ is a homomorphism on H; this simply amounts to solving the system of linear equations

$$\sigma(u_j) = m_{j1}x_1 + \cdots + m_{jp}x_p$$

for some rational numbers x_k. If there is a linear dependence among the right-hand members of these equations, say $\sum_j n_j m_{jk} = 0$ for $1 \leqq k \leqq p$, and some integers n_j, then

$$\prod_j u_j^{n_j} \in \mathrm{P} \cap [\mathrm{H}, \mathrm{H}],$$

and consequently $\sum_j n_j \sigma(u_j) = \sigma(\prod_j u_j^{n_j}) = 0$. These equations are thereby consistent, and do admit solutions.

Definition. The *character class* (associated to the presentation $\Gamma \cong \mathrm{H}/\mathrm{P}$) of a factor of automorphy ν is the element

$$\chi_1(\nu) = \phi(\hat{\sigma}) \in \mathrm{Hom}(\mathrm{P} \cap [\mathrm{H}, \mathrm{H}]/[\mathrm{P}, \mathrm{H}]; Z),$$

where $\hat{\sigma}$ is any homomorphism of the form (5) and ϕ is the canonical isomorphism of Lemma 2. As a consequence of (6), the character class is uniquely determined by the factor of automorphy ν alone. If μ and ν are two factors of automorphy, then $\chi_1(\mu\nu) = \chi_1(\mu) + \chi_1(\nu)$; thus χ_1 is a homomorphism of the group of factors of automorphy into $\mathrm{Hom}(\mathrm{P} \cap [\mathrm{H}, \mathrm{H}]/[\mathrm{P}, \mathrm{H}]; Z)$. The extent to which the character class is independent of the presentation $\Gamma \cong \mathrm{H}/\mathrm{P}$ will be discussed in the supervening section, in which the underlying algebraic structure will be examined.

LEMMA 3. *Two sets of factors of automorphy μ and ν have the same character class if and only if there exists a character $\hat{c} = \{c_T\}$ of the group Γ and a summand of automorphy σ such that $\mu_T(z) = c_T \nu_T(z) \exp \sigma_T(z)$.*

Proof. Introducing the quotient factor $\eta = \mu/\nu$, the factors μ and ν have the same character class if and only if $\chi_1(\eta) = 0$. Corresponding to the presentation $\Gamma \cong \mathrm{H}/\mathrm{P}$, select holomorphic functions $\tau_t(z)$ for each $t \in \mathrm{H}$ such that $\tau_{st}(z) = \tau_s(Tz) + \tau_t(z)$ and that $\eta_T(z) = \exp \tau_t(z)$; this may be done in the manner in which we defined the character class, for example. Any character $\hat{c} = \{c_T\}$ of the group Γ can be written in the form $c_T = \exp 2\pi i\, a_t$, where $a_t = \hat{a}(t)$ and

$$\hat{a} \in \mathrm{Hom}(\mathrm{P}/[\mathrm{P}, \mathrm{H}], \mathrm{H}/[\mathrm{P}, \mathrm{H}]; Z, \bar{Q}) \tag{7}$$

for $\bar{Q}$ the additive group of real numbers. Letting $\sigma_t(z) = \tau_t(z) - 2\pi i a_t$, it follows that

$$\mu_T(z)/\nu_T(z) = \eta_T(z) = c_T \exp \sigma_t(z).$$

The functions $\sigma_t(z)$ will define a summand of automorphy for a suitable choice of $\hat{a}$ if and only if $\sigma_t(z) \equiv 0$ for all $t \in \mathrm{P}$, or what is the same, if and only if $\hat{\sigma} = 0$ for the homomorphism $\hat{\sigma}$ defined by (5). Since $\hat{\sigma} = \hat{\tau} - \hat{a}$, it is possible to select $\hat{a}$ such that $\hat{\sigma} = 0$ if and only if $\hat{\tau}$ belongs to the group (7) containing all possible $\hat{a}$; by Lemma 2, this in turn is equivalent to the fact that $\phi(\hat{\tau}) = \chi_1(\eta) = 0$.

THEOREM 2. *Two factors of automorphy are equivalent if and only if they have the same character class.*

Proof. By Lemma 3, factors μ and ν have the same character class if and only if there is a character $\hat{c} = \{c_T\}$ of Γ and a summand of automorphy σ such that $\mu_T(z) = c_T \nu_T(z) \exp \sigma_T(z)$. By Theorem 1, there exists for any summand σ a holomorphic function $g(z)$ such that $g(Tz) = g(z) + \sigma_T(z) + 2\pi i a_T$ for some real numbers a_T. Letting $b_T = c_T \exp[-2\pi i a_T]$ and $h(z) = \exp g(z)$, μ and ν have the same character class if and only if $\mu_T(z) = b_T \nu_T(z) h(Tz)/h(z)$, which was to be proved.

The final cohomological classification of factors of automorphy follows trivially. From each class of equivalent factors of automorphy select a basic element ν; any other factor in the same equivalence class can be written in the form $\mu_T(z) = c_T \nu_T(z) \exp \sigma_T(z)$ for some summand σ. The mapping $T \to c_T$ is actually a homomorphism on the abelianized group $\Gamma/[\Gamma, \Gamma]$; one then sees trivially that c_T and σ may be modified in such a manner that $c_T = 1$ except on the torsion elements of $\Gamma/[\Gamma, \Gamma]$. The homomorphism $T \to c_T$ is then unique, and will be called the torsion class $T_\nu(\mu)$ of μ with respect to ν; the class $\chi_2(\sigma) = \chi_{2,\nu}(\mu)$ of the summand σ may also be considered as an invariant associated to the factor μ with respect to ν. The three invariants $\chi_1(\mu)$, $T_\nu(\mu)$, and $\chi_{2,\nu}(\mu)$ then completely characterize the cohomology classes of factors.

5. The preceding analysis clearly demonstrated that the equivalence classification of factors of automorphy is an algebraic consequence of Theorem 1. As a short digression, we shall rephrase the above discussion in a more abstract setting which emphasizes these formal aspects of the argument. Let Γ be a group, which for the sake of convenience we shall continue to assume

finitely generated, and G be its group ring $Z(\Gamma)$. Consider the following collection of G-modules and G-homomorphisms:

$$\begin{tikzcd}
0 \arrow[r] & J \arrow[r, "j"] & A \arrow[r, "\phi"] & M \arrow[r] & 0 \\
0 \arrow[r] & J \arrow[r, "k"] \arrow[u, "i"] & P \arrow[r, "\psi"] \arrow[u, "\alpha"] & R \arrow[r] \arrow[u, "\beta"] & 0
\end{tikzcd}$$

The two horizontal rows are exact sequences, the diagram is commutative, and i is the identity map. From this one derives the diagram:

$$\begin{tikzcd}
\cdots \arrow[r] & H^1(\Gamma, J) \arrow[r, "j^*"] & H^1(\Gamma, A) \arrow[r, "\phi^*"] & H^1(\Gamma, M) \arrow[r, "\delta_1^*"] & H^2(\Gamma, J) \arrow[r] & \cdots \\
\cdots \arrow[r] & H^1(\Gamma, J) \arrow[r, "k^*"] \arrow[u, "i^*"] & H^1(\Gamma, P) \arrow[r, "\psi^*"] \arrow[u, "\alpha^*"] & H^1(\Gamma, R) \arrow[r, "\delta_2^*"] \arrow[u, "\beta^*"] & H^2(\Gamma, J) \arrow[r] \arrow[u, "i^*"] & \cdots
\end{tikzcd}$$

The two horizontal rows are again exact sequences, the diagram is commutative, and i^* is the identity map. It follows immediately from this diagram that whenever α^* is an isomorphism onto, then β^* is an isomorphism which has as image the subgroup $\delta_1^{*-1}i^*\delta_2^*[H^1(\Gamma, J)] \subset H^1(\Gamma, M)$. That is, if we define a homomorphism $\chi: H^1(\Gamma, M) \to H^2(\Gamma, J)/i^*\delta_2^*[H^1(\Gamma, R)]$ by associating to each $m \in H^1(\Gamma, M)$ the coset containing $\delta_1^*(m)$, then the image of β^* is the kernel of χ. Two 1-cocycles μ and ν of the group Γ with coefficients in M are therefore cohomologous to the image under β^* of a 1-cocycle c having coefficients in R if and only if $\chi(\mu) = \chi(\nu)$; the 1-cocycle c is unique when it exists.

Returning to our particular case, set $J = Z$, $A = \mathfrak{A}(D)$, $M = \mathfrak{M}(D)$, $P =$ additive group of real numbers, $R =$ one-dimensional unitary group, j and $k =$ injection maps, and ϕ and $\psi =$ maps: $x \to \exp 2\pi i x$. Theorem 1 states that α^* is an isomorphism onto; hence two factors μ and ν are equivalent if and only if $\chi(\mu) = \chi(\nu)$. Applying Lemma 2, one sees that

$$H^2(\Gamma, Z)/i^*\delta_2^*[H^1(\Gamma, R)] \cong \mathrm{Hom}(\mathrm{P} \cap [\mathrm{H}, \mathrm{H}]/[\mathrm{P}, \mathrm{H}]; Z)$$

whenever Γ is presented as the factor group of a finitely generated free group H modulo a group of relations P, and thus identifies χ under this isomorphism with the character class.

III. The Rôle of the Character Class.

6. A *relatively automorphic function* associated to a factor of automorphy ν is a meromorphic function $f(z)$ on D such that $f(Tz)=\nu_T(z)f(z)$. The set of all relatively automorphic functions associated to the factor ν form a complex linear space $\mathcal{L}(\nu)$. If μ is a factor of automorphy cohomologous to ν, there exists a function $h(z)$ holomorphic and nowhere-vanishing on D such that $h(Tz)=h(z)\mu_T(z)/\nu_T(z)$; then whenever $f(z)\in\mathcal{L}(\nu)$, the function $T_{\mu\nu}f(z)=h(z)f(z)\in\mathcal{L}(\mu)$. The mapping $T_{\mu\nu}:\mathcal{L}(\nu)\rightarrow\mathcal{L}(\mu)$ so defined is an *analytic isomorphism*; that is, in addition to being an isomorphism of the complex linear spaces, it satisfies the following two conditions: (i) $T_{\mu\nu}$ preserves the structure of the zeros and poles of the functions involved, in the sense that $T_{\mu\nu}f(z)/f(z)$ is holomorphic and nowhere-vanishing, (ii) $T_{\mu\nu}$ preserves the representation of automorphic functions, in the sense that an automorphic function represented as $f(z)=f_1(z)/f_2(z)$ for $f_j(z)\in\mathcal{L}(\nu)$ is also represented as $f(z)=T_{\mu\nu}f_1(z)/T_{\mu\nu}f_2(z)$. Conversely it is clear that whenever $\mathcal{L}(\mu)$ and $\mathcal{L}(\nu)$ are non-vacuous and analytically isomorphic, the factors μ and ν are cohomologous. This expresses the function-theoretic significance of the concept of cohomologous factors of automorphy.

A singular 1-simplex σ^1 of D is in *general position* with respect to a function $f(z)$ meromorphic on D if $f(z)$ is finite-valued and non-zero at each point of the support of σ^1; a singular 1-chain κ^1 of D is in general position with respect to $f(z)$ if each component simplex in κ^1 is in general position. The set of all singular 2-chains κ^2 of D whose boundaries are in general position with respect to $f(z)$ form an abelian group, and $f(z)$ defines a linear function on this group by

$$c(f)[\kappa^2]=\frac{1}{2\pi i}\int_{\partial\kappa^2}\frac{df(z)}{f(z)}.$$

THEOREM 3. *Let χ_1 be the character class associated to a presentation $\Gamma\cong\mathrm{H}/\mathrm{P}$ and Φ be the homomorphism* (2). *Then for any factor of automorphy ν and associated relatively automorphic function $f(z)$,*

$$c(f)[\Phi(v)]=\chi_1(\nu)[v] \tag{8}$$

for every $v\in\mathrm{P}\cap[\mathrm{H},\mathrm{H}]$.

Proof. Reverting to the notation of Section 1, it is clear that to each $v\in\mathrm{P}\cap[\mathrm{H},\mathrm{H}]$ we may associate one of the standard singular 2-chains $\kappa^2(v)$

whose boundary $\kappa^1(v) = \partial\kappa^2(v)$ is in general position with respect to $f(z)$; the precise meaning of (8) is then

$$c(f)[\kappa^2(v)] = \chi_1(\nu)[v]$$

whenever $\kappa^2(v)$ is in general position with respect to $f(z)$. It follows from the definition of $\kappa^1(v)$ that

$$\begin{aligned}\int_{\kappa^1(v)} df(z)/f(z) &= \sum_r \int_{\sigma_r^1 - \bar{\sigma}_r^1} df(z)/f(z)\\ &= -\sum_r \int_{\sigma_r^1} d\nu[T_{\alpha(r)}{}^{\epsilon(r)}](z)/\nu[T_{\alpha(r)}{}^{\epsilon(r)}](z),\end{aligned}$$

where we have written $\nu[T](z)$ for $\nu_T(z)$. Introducing functions

$$\sigma_r(z) = \log \nu[T_{\alpha(r)}{}^{\epsilon(r)}](z)$$

one secures

$$\begin{aligned}\int_{\kappa^1(v)} df(z)/f(z) &= -\sum_r [\sigma_r(T_{\alpha(j(r))}{}^{\epsilon(j(r))} \cdots T_{\alpha(1)}{}^{\epsilon(1)} z_0)\\ &\qquad - \sigma_r(T_{\alpha(r-1)}{}^{\epsilon(r-1)} \cdots T_{\alpha(1)}{}^{\epsilon(1)} z_0)]\\ &= \sigma_v(z_0) = 2\pi i \chi_1(\nu)[v].\end{aligned}$$

Thus

$$c(f)[\kappa^2(v)] = \frac{1}{2\pi i}\int_{\kappa^1(v)} df(z)/f(z) = \chi_1(\nu)[v], \tag{9}$$

which was to be proved.

For an interpretation of this theorem in slightly different terms, let us recall that a *divisor* $\mathfrak{D}$ on the manifold D is a finite formal sum $\mathfrak{D} = \sum_j m_j V_j$, where V_j are irreducible $(n-1)$-dimensional complex analytic subvarieties of D and m_j are positive or negative integers. In the open sets U_λ of a sufficiently fine covering of D, each subvariety V_j may be represented as the locus of the zeros of a function $d_{j\lambda}(z)$ which has no multiple factors at any point of U_λ; the function $d_\lambda(z) = \prod_j d_{j\lambda}(z)^{m_j}$ is called a minimal local equation of the divisor $\mathfrak{D}$ in U_λ. If $f(z)$ is a meromorphic function on D, there is a unique divisor $\mathfrak{D}(f)$ having $f(z)$ as a minimal local equation at each point; $\mathfrak{D}(f)$ is called the divisor of the function $f(z)$. A divisor $\mathfrak{D}$ with minimal local equations $d_\lambda(z)$ is Γ-invariant if $d_\lambda(Tz)/d_\mu(z)$ is holomorphic and non-vanishing in $U_\lambda \cap T^{-1}U_\mu$, for every $T \in \Gamma$. Clearly $\mathfrak{D}(f)$ is Γ-invariant whenever $f(z)$ is a relatively automorphic function; conversely, on those manifolds D with the property that every divisor is the divisor of a

global meromorphic function, every Γ-invariant divisor is the divisor of a relatively automorphic function. The manifolds D of this type include the manifolds of the particular examples considered in Chapter IV. This characterization of relatively automorphic functions also illustrates their role in the analytic representation of automorphic functions; for such manifolds D, every automorphic function can be represented as the quotient of two holomorphic relatively automorphic functions.

A point on an irreducible analytic subvariety V is called a simple point if there are local coordinates $z_1, \cdots z_n$ centered at the point, in terms of which V is locally a plane section $z_{p+1} = \cdots = z_n = 0$. The points which are not simple lie in a closed subvariety of V which is of still lower dimension, and may be called the singular set of V. Thus outside of its singular set, V may be given the structure of a complex p-dimensional manifold by introducing local coordinates $z_1, \cdots, z_p$ in the appropriate coordinate neighborhoods; the analytic structure induces a definite orientation on V in the usual manner. A (differentiable) singular 2-simplex σ^2 of D is in general position with respect to a divisor $\mathcal{D} = \sum_j m_j V_j$ if the support of σ^2 meets the point set $\bigcup_j V_j$ in finitely many interor points of σ^2, each of which is a simple point lying on but one of the subvarieties V_j, and at each of which the simplex σ^2 and set V_j meet transversally in the sense that their tangent spaces generate a full $2n$-dimensional space. A singular 2-chain κ^2 is in general position with respect to $\mathcal{D}$ if each of its simplices is. The set of all singular 2-chains in general position with respect to $\mathcal{D}$ form an abelian group, and $\mathcal{D}$ defines a linear function on this group by introducing the sum of the intersection multiplicities K. I. $(\mathcal{D}, \kappa^2)$, just as in the definition of the Kronecker index. It is clear from the Cauchy residue formula in one variable that whenever $\mathcal{D}$ is the divisor of a meromorphic function $f(z)$ on D, K. I. $(\mathcal{D}(f), \kappa^2) = c(f)[\kappa^2]$. Hence by Theorem 3 it follows that for any $f(z) \in \mathcal{L}(\nu)$,

$$\text{K. I.}\,(\mathcal{D}(f), \Phi(v)) = \chi_1(\nu)[v].$$

Now suppose that Γ contains no transformations with fixed points. For each singular 2-chain $\kappa^2{}_\Gamma$ on D/Γ select a 2-chain κ^2 on D such that $\rho(\kappa^2) = \kappa^2{}_\Gamma$; the fact that this is possible follows from the covering homotopy theorem for example. Then $\mathcal{D}(f)$ defines a linear function on the group of all 2-chains $\kappa^2{}_\Gamma$ of D/Γ which are in general position with respect to $\rho(\mathcal{D}(f))$ by K. I.$_\Gamma(\mathcal{D}(f), \kappa^2{}_\Gamma) =$ K. I. $(\mathcal{D}(f), \kappa^2)$, since this is clearly independent of the choice of κ^2. Let us call this function the 2-cocycle dual to

the divisor $\mathcal{D}(f)$. The map Φ of Section 1 is an isomorphism onto, and hence induces an onto isomorphism

$$\Phi^{*}: \mathrm{Hom}(H_2(D/\Gamma)/S_2(D/\Gamma); Z) \to \mathrm{Hom}(\mathrm{P} \cap [\mathrm{H}, \mathrm{H}]/[\mathrm{P}, \mathrm{H}]; Z).$$

Since $H^2(D/\Gamma; Z) \cong \mathrm{Hom}(H_2(D/\Gamma); Z)$, the character class of any factor of automorphy represents a 2-dimensional integral cohomology class of D/Γ, which is aspherical in the sense that it vanishes on every spherical cycle. Theorem 3 asserts that the 2-cocycle dual to the divisor of a relatively automorphic function associated to a factor of automorphy ν represents the two-dimensional cohomology class defined by the character class of ν. Further, the divisor of any relatively automorphic function is aspherical in the above sense.

7. We shall continue to assume in this section that the group Γ contains no transformations with fixed points. Thus D/Γ is itself a complex manifold, and for the divisor $\mathcal{D}(f)$ of any relatively automorphic function $f(z) \in \mathcal{L}(\nu)$, $\rho(\mathcal{D}(f))$ is a divisor on the manifold D/Γ. In order to discuss the homological properties of divisors further, we shall assume that the divisors $\rho(\mathcal{D}(f)) = \sum_j m_j\rho(V_j)$, where the subvarieties V_j are given their natural orientations, can be expressed as singular cycles of D/Γ of class C^r $(r \geqq 1)$. This is a considerably weakened form of the strong covering theorem for analytic manifolds, which asserts that the manifold D/Γ can be covered by a simplicial complex of class C^r for arbitrary r in such a manner that $\rho(\mathcal{D}(f))$ is a subcomplex. For further discussion of this theorem, see for example [15, 16].

It should be pointed out here that the 2-cocycle dual to the divisor $\mathcal{D}(f)$, as defined previously, is just the cocycle dual to the singular cycle $\rho(\mathcal{D}(f))$ in the usual singular sense. Utilizing the singular form of the de Rham representation [12], for any $(2n-2)$-dimensional cohomology class ϕ on D/Γ, represented by a Γ-invariant differential form $\phi(z)$ on D, $\phi[\mathcal{D}] = \sum_j m_j \int_{V_j} \phi(z)$, the integration being extended over the C^1 singular cycles $\rho(V_j)$ on D/Γ. Further, letting $\chi_1(\nu)$ also denote the 2-dimensional cohomology class defined by the character class of the factor of automorphy ν and dual to $\mathcal{D}(f)$, $\phi[\mathcal{D}] = (\phi \cup \chi_1(\nu))[D/\Gamma]$, where $[D/\Gamma]$ is the fundamental cycle of the manifold D/Γ.

LEMMA 4. *Let W be an analytic subvariety of D/Γ of complex dimension*

$n-1$, and $\theta(z)$ be a Γ-invariant differential form of type $(n-1,0)$ on D with non-trivial support. Then

$$i^{(1-n^2)}\int_W \theta(z)\wedge\bar{\theta}(z)>0.$$

Proof. Consider firstly a singular simplex σ^{2n-2} of W which is disjoint from the singular set of W and contains an open set in the support of $\theta(z)$. Then complex coordinates $z_1,\cdots,z_n$, or alternatively the real coordinates $x_1,\cdots,x_{2n}$ with $z_j=x_{2j-1}+ix_{2j}$, can be introduced in an open neighborhood of σ^{2n-2} in D/Γ in such a manner that W is locally the subvariety $z_n=0$. Writing the differential form $\theta_n(z)$ as

$$\theta(z)=\sum_{j=1}^{n}\theta_{1\cdots j-1\,j+1\cdots n}(z)\,dz^1\wedge\cdots\wedge dz^{j-1}\wedge dz^{j+1}\wedge\cdots\wedge dz^n,$$

we have

$$\begin{aligned}\int_{\sigma^2}&\theta(z)\wedge\bar{\theta}(z)\\ &=\int_{\sigma^2}\theta_{1\cdots n-1}(z)\,\bar{\theta}_{1\cdots n-1}(z)\,dz^1\wedge\cdots\wedge dz^{n-1}\wedge d\bar{z}^1\wedge\cdots\wedge d\bar{z}^{n-1}\\ &=2^{n-1}i^{(n-1)(n+1)}\int_{\sigma^2}|\,\theta_{1\cdots n-1}(x)\,|^2\,dx^1\wedge\cdots\wedge dx^{2n}.\end{aligned}$$

Thus

$$i^{(1-n^2)}\int_{\sigma^2}\theta(z)\wedge\bar{\theta}(z)>\epsilon>0.$$

Select a barycentric subdivision of the chain W which is sufficiently fine that the integral of $i^{(1-n^2)}\theta(z)\wedge\theta(z)$ over all simplices of W which meet the singular set of W is less than ϵ in absolute value. The assertion of the lemma follows immediately.

We shall now derive from this lemma a property of the character class of factors of automorphy which admit holomorphic relatively automorphic functions, generalizing a well-known theorem of Frobenius on complex tori [9]. A factor of automorphy ν is called *positive* if it admits an associated relatively automorphic function $f(z)$ which is holomorphic on D. Let $\{\theta_\alpha\}$ be a basis for the subgroup of all complex-valued cohomology classes of dimension $n-1$ on D/Γ which can be represented in the sense of de Rham by differential forms $\{\theta_\alpha(z)\}$ of type $(n-1,0)$, and $\{\bar{\theta}_\alpha\}$ be the complex conjugate classes represented by the differential forms $\{\bar{\theta}_\alpha(z)\}$.

THEOREM 4. *Letting ν be a positive factor of automorphy with character*

class $\chi_1(\nu)$ *considered as a 2-cocycle of* D/Γ, *and* $[D/\Gamma]$ *be the fundamental cycle of the manifold* D/Γ, *the matrix*

$$i^{(1-n^2)}((\chi_1(\nu)\cup\theta_\alpha\cup\bar{\theta}_\beta)[D/\Gamma])_{\alpha\beta}$$

is positive Hermitian. (Γ is assumed to contain no transformations with fixed points.)

Proof. Let $f(z)\in\mathcal{L}(\nu)$ be a holomorphic relatively automorphic function associated to the factor ν, and $\mathcal{D}(f)=\sum_j m_jV_j$ be its divisor; then since $f(z)$ is holomorphic, $m_j>0$ for all j. For any arbitrary complex constants $\{\xi_\alpha\}$, not all of which are zero, $\theta(z)=\sum_\alpha\xi_\alpha\theta_\alpha(z)$ will be a Γ-invariant differential form of type $(n-1,0)$ on D which represents the cohomology class $\sum_\alpha\xi_\alpha\theta_\alpha$. Now on the one hand by Lemma 4,

$$\sum_j i^{(1-n^2)}m_j\int_{V_j}\theta(z)\wedge\bar{\theta}(z)>0.$$

On the other hand

$$\begin{aligned}\sum_j i^{(1-n^2)}m_j\int_{V_j}\theta(z)\wedge\bar{\theta}(z)\\ &=\sum_j\sum_{\alpha,\beta}i^{(1-n^2)}m_j\xi_\alpha\bar{\xi}_\beta\int_{V_j}\theta_\alpha(z)\wedge\bar{\theta}_\beta(z)\\ &=\sum_{\alpha,\beta}i^{(1-n^2)}\xi_\alpha\bar{\xi}_\beta(\theta_\alpha\cup\bar{\theta}_\beta)[\mathcal{D}(f)]\\ &=\sum_{\alpha,\beta}i^{(1-n^2)}\xi_\alpha\bar{\xi}_\beta(\chi_1(\nu)\cup\theta_\alpha\cup\bar{\theta}_\beta)[D/\Gamma].\end{aligned}$$

This demonstrates the theorem.

8. The method of Section 6 yields a criterion for determining which elements of $\mathrm{Hom}(\mathrm{P}\cap[\mathrm{H},\mathrm{H}]/[\mathrm{P},\mathrm{H}];Z)$ arise as possible character classes of factors of automorphy. If $g(z)$ is any C^∞ function on the manifold D for which $\bar{\partial}\partial g(z)$ is a Γ-invariant differential form, associate to $g(z)$ the element $\chi_g\in\mathrm{Hom}(\mathrm{P}\cap[\mathrm{H},\mathrm{H}]/[\mathrm{P},\mathrm{H}];C^+)$ defined by

$$\chi_g(v)=\int_{\kappa^2(v)}\bar{\partial}\partial g(z).$$

Here C^+ is the additive group of complex numbers and $\kappa^2(v)$ is one of the standard 2-cycles representing the image $\Phi(v)$ of the homomorphism Φ of Section 1; the element $\chi_g(v)$ is clearly well-defined.

THEOREM 5. *If $g(z)$ is a C^∞ function on D such that $\bar\partial\partial g(z)$ is Γ-invariant and χ_g is integral-valued, then some multiple of χ_g is the character class of a factor of automorphy; conversely the character class of any factor of automorphy can be so represented.*

Proof. To demonstrate the converse first, note that in the proof of Theorem 3 the second equality in (9) remains true when $df(z)/f(z)$ is replaced by any closed differential form $\rho(z)$ of degree 1 which is well-defined in a neighborhood of $\kappa^1(v)$ and satisfies $\phi(Tz) = \phi(z) + d\log \nu_T(z)$ in that neighborhood. Applying Lemma 1, select a C^∞ function $g(z)$ which is real-valued and satisfies $g(Tz) = g(z) + \frac{1}{2\pi i}\log|\nu_T(z)|^2$, where the real values of the logarithm are selected; then $\partial g(Tz) = \partial g(z) + \frac{1}{2\pi i} d\log \nu_T(z)$. Then by (9) and Stokes' theorem,

$$\chi_1(\nu)[v] = \int_{\kappa^1(v)} \partial g(z) = \int_{s^2(v)} \bar\partial\partial g(z) = \chi_g(v)$$

for every $v \in \mathrm{P} \cap [\mathrm{H}, \mathrm{H}]$.

If $\partial\partial g(z)$ is a Γ-invariant differential form on D, then for any transformation $T \in \Gamma$ consider the form $\theta_T(z) = \partial g(Tz) - \partial g(z)$. Since $\bar\partial\theta_T(z) = \bar\partial\partial g(Tz) - \bar\partial\partial g(z) = 0$, $\theta_T(z)$ are closed, holomorphic differential forms, and it follows from their definition that $\theta_{ST}(z) = \theta_S(Tz) + \theta_S(z)$. For each free generator t_j of H select an indefinite integral $\sigma_{t_j}(z) = \int^z \theta_{T_j}(\zeta)$, and extend the functions so constructed to a collection indexed by all $t \in \mathrm{H}$ and satisfying $\sigma_{st}(z) = \sigma_s(Tz) + \sigma_t(z)$. Whenever $r \in \mathrm{P}$, $\hat\sigma(r) = \sigma_r(z)$ is a constant, and for all $s \in \mathrm{H}$, $\hat\sigma(srs^{-1}) = \hat\sigma(r)$. It is again clear from (9) that $\hat\sigma(v) = \chi_g(v)$ for all $v \in \mathrm{P} \cap [\mathrm{H}, \mathrm{H}]$. Thus if $\hat\sigma(r) \in Z$ for all $r \in \mathrm{P}$, then $\nu_T(z) = \exp \sigma_t(z)$ defines a factor of automorphy with character class $\chi_1(\nu) = \chi_g$.

Even when $\hat\sigma(r)$ are not always integers, it is still true by hypothesis that $\hat\sigma(v) = \chi_g(v) \in Z$ for all $v \in \mathrm{P} \cap [\mathrm{H}, \mathrm{H}]$. Moreover, applying Lemma 2 after replacing H by P and P by $\mathrm{P} \cap [\mathrm{H}, \mathrm{H}]$, there is a homomorphism α of P into the additive group of rational numbers such that $\hat\alpha(v) = \hat\sigma(v)$ for all $v \in \mathrm{P} \cap [\mathrm{H}, \mathrm{H}]$. Let m be the least common multiple of the denominators of the rational numbers $\hat\alpha(r_j)$ for some finite set of free generators r_j of P, and let $\tau(r) = m\hat\sigma(r) - m\hat\alpha(r)$ for $r \in \mathrm{P}$. By Lemma 2 again, $\hat\tau(r)$ may be extended to a homomorphism of H into the rationals, which we shall also denote by $\hat\tau$. The functions $\tau_t(z) = m\sigma_t(z) - \hat\tau(t)$ have the same values as $m\sigma_t(z)$ whenever $t \in \mathrm{P} \cap [\mathrm{H}, \mathrm{H}]$, and $\tau_t(z) \in Z$ whenever $t \in \mathrm{P}$. Thus $\nu_T(z) = \exp \tau_t(z)$ defines a factor of automorphy with the character class $\chi_1(\nu) = m\chi_g$.

We shall call a group Γ *integral* if every form χ_g itself represents the character class of a factor of automorphy. The examples to be considered in Chapter 4 will clearly be seen to be integral. It should be noted here that, upon applying the decomposition theorems for the operators ∂ and $\bar{\partial}$ in the obvious manner, $\bar{\partial}\partial g(z) = \psi(z) + \bar{\partial}\partial h(z)$ where $\psi(z)$ is a harmonic differential form and $h(z)$ is a Γ-invariant C^∞ function on D. Therefore for the purposes of Theorem 5 it is sufficient to consider only those functions $g(z)$ for which $\bar{\partial}\partial g(z)$ is a harmonic differential form.

Now assuming that the group Γ contains no transformations with fixed points, the differential forms $\phi(z) = \bar{\partial}\partial g(z)$ of Theorem 5 represent in the de Rham sense the character class $\chi_1(\nu)$ considered as a 2-cocycle of D/Γ. Theorem 4 may be expressed more analytically in terms of this differential form.

COROLLARY *to Theorem* 4. *If* $\phi(z) = \sum_{j,k} \phi_{j\bar{k}}(z) dz^j \wedge d\bar{z}^k$ *is a differential form representing the character class* $\chi_1(\nu)$ *of a positive factor of automorphy, then the Hermitian matrix* $(-i\phi_{j\bar{k}}(z))$ *is positive definite in the mean on harmonic forms, in the following sense: for any non-trivial closed differential form*

$$\theta(z) = \sum_j (-1)^j \theta_j(z) dz^1 \wedge \cdots \wedge dz^{j-1} \wedge dz^{j+1} \wedge \cdots \wedge dz^n,$$

$$-i \int_{D/\Gamma} \sum_{j,k} \phi_{j\bar{k}}(z) \theta_j(z) \bar{\theta}_k(z) dv > 0,$$

where dv is the positive volume element on the manifold D/Γ. (Γ is assumed to contain no transformations with fixed points.)

Proof. Theorem 4 asserts that

$$i^{(1-n^2)}(\chi_1(\nu) \cup \theta \cup \bar{\theta})[D/\Gamma] > 0,$$

where θ is the cohomology class represented by the differential form $\theta(z)$. However in terms of differential forms

$$(\chi_1(\nu) \cup \theta \cup \bar{\theta})[D/\Gamma] = \int_{D/\Gamma} \phi(z) \wedge \theta(z) \wedge \bar{\theta}(z)$$

$$= 2^n i^{(n^2-2)} \int_{D/\Gamma} \sum_{j,k} \phi_{j\bar{k}}(z) \theta_j(z) \bar{\theta}_k(z) dx^1 \wedge \cdots \wedge dx^{2n},$$

where $z_j = x_{2j-1} + ix_{2j}$, from which the assertion follows.

9. It is perhaps appropriate to discuss at this point the connection between factors of automorphy as considered here and the related concept of

a complex line bundle [14]. If Γ has no fixed points, then every factor of automorphy defines a complex line bundle on the manifold D/Γ in the obvious manner; the presence of fixed points means not only that D/Γ need not be a complex manifold, but also that the line bundle defined may be locally multiple-valued. To proceed in the opposite direction, any complex line bundle on D/Γ, when D/Γ is a complex manifold, induces a complex line bundle on D. If D is a Stein manifold, for example [5, 6], then all topologically trivial line bundles on D are also analytically trivial, and are thus equivalent to line bundles induced by factors of automorphy; in particular if $H^2(D, R) = 0$ for a Stein manifold D, all line bundles on D/Γ are equivalent to bundles induced by factors of automorphy.

In the present note, factors of automorphy have been treated on their own, without utilizing their relationship to complex line bundles. This method has the advantages that the concepts and constructions can be developed in terms of the group Γ and its action on the manifold D, and that they express in a natural manner the global nature and multiplicity at fixed points which distinguish factors of automorphy from line bundles. In addition, the forms in which factors of automorphy appear, and such results as Theorem 4, are of more interest from a function-theoretic point of view than from the point of view of line bundles.

IV. Examples.

10. Our first example illustrates a method for constructing non-trivial factors of automorphy in fairly general cases. Letting $\{\omega_\alpha(z)\}$ be a basis for the complex linear space of abelian differentials on D/Γ and $\{w_\alpha(z)\}$ be the corresponding abelian integrals, introduce the functions $g(z) = \sum_{\alpha,\beta} \xi_{\alpha\beta} w_\alpha(z)\bar{w}_\beta(z)$ for arbitrary complex constants $\xi_{\alpha\beta}$. The differential forms

$$(10) \qquad \bar{\partial}\partial g(z) = -\sum_{\alpha,\beta} \xi_{\alpha\beta}\omega_\alpha(z) \wedge \bar{\omega}_\beta(z)$$

are obviously Γ-invariant, so that they define elements

$$\chi_g \in \mathrm{Hom}(\mathrm{P} \cap [\mathrm{H}, \mathrm{H}]/[\mathrm{P}, \mathrm{H}]; C^+)$$

as in Section 8. If χ_g is integral valued for some choice of $\xi_{\alpha\beta}$, Theorem 5 asserts that a multiple of χ_g represents the character class of a factor of automorphy. The proof of Theorem 5 was actually constructive in nature; examining that proof, one sees that the factors of automorphy so represented are given explicitly by

$$(11) \qquad \nu_T(z) = \exp[\sum_{\alpha\beta} \xi_{\alpha\beta}\bar{\omega}_\beta(T) w_\alpha(z) + \zeta_T],$$

or by integral powers of these functions if necessary; the ζ_T are suitably chosen complex constants and $\omega_\beta(T)$ are the periods of the abelian integrals $w_\beta(z)$. Factors of automorphy of the form (11) may be called generalized theta factors, and the associated relatively automorphic functions generalized theta functions [18].

It is perhaps of some interest to examine the classical theta functions [19, 20] from this point of view. For this purpose, let D be the entire n-dimensional complex affine space and Γ be a group of translations generated by the translations along $2n$ real linearly independent vectors in D. The abelian differentials are then simply the differential forms dz^α, $\alpha = 1, \cdots, n$, and the periods of the associated abelian integrals corresponding to a transformation $T \in \Gamma$ are the components of the vector representing the translation T. A differential form $\phi = \sum_{\alpha,\bar\beta} \phi_{\alpha\bar\beta}(z)\, dz^\alpha \wedge d\bar z^\beta$ is harmonic and Γ-invariant if and only if the coefficients $\phi_{\alpha\bar\beta}(z)$ are harmonic, Γ-invariant functions on D, hence constants; that is to say, all harmonic Γ-invariant differential forms of type $(1,1)$ are of the form (10). Therefore all factors of automorphy are equivalent to factors (11), which in this case have the even simpler form

$$\nu_T(z) = \exp\big[\sum_{\alpha,\bar\beta} \xi_{\alpha\bar\beta}\bar\omega^\beta(T) z_\alpha + \zeta_T\big]. \tag{12}$$

The latter statement was first proved in more than one complex variable by Appell [1].

The character class has a well-known representation in this case. Letting $T_1, \cdots, T_{2n}$ be the translations generating the group Γ, we may write $\Gamma \cong \mathrm{H}/\mathrm{P}$ where H is a free group on corresponding generators $t_1, \cdots, t_{2n}$ and P is the normal subgroup of H generated by the words $v_{jk} = t_j t_k t_j^{-1} t_k^{-1}$. Since $\mathrm{P} \cap [\mathrm{H}, \mathrm{H}] = \mathrm{P}$, the group Γ is certainly integral. The character class $\chi_1(\nu)$ of the factor (12) is determined by the values

$$\chi_1(\nu)[v_{jk}] = \frac{1}{2\pi i} \sum_{\alpha,\bar\beta} \xi_{\alpha\bar\beta}[-\omega_{\alpha j}\bar\omega_{\beta k} + \omega_{\alpha k}\bar\omega_{\beta j}].$$

In matrix form $\Omega = (\omega_{\alpha j})$ is the period matrix, either in the classical sense or in the sense of Section 2, $\Xi = (\xi_{\alpha\bar\beta})$, and $\mathrm{X}_1(\nu) = (\chi_1(\nu)[v_{jk}])$ is given by

$$\mathrm{X}_1(\nu) = \frac{1}{2\pi i}[{}^t\bar\Omega\, {}^t\Xi\, \Omega - {}^t\Omega\, \Xi\, \bar\Omega]. \tag{13}$$

The differential form $\sum_{\alpha,\bar\beta} \xi_{\alpha\bar\beta}\, dz^\alpha \wedge d\bar z^\beta$ therefore represents the character class of a factor of automorphy if and only if the matrix (13) is integral; furthermore a group Γ with period matrix Ω will admit a non-trivial factor of auto-

morphy if and only if there exists a constant matrix Ξ such that (13) is integral.

The above is one part of the condition that Ω be a Riemann matrix. If we require that the group Γ with Ω as period matrix admit a positive factor of automorphy, we secure, as with Frobenius [9], the complete condition that Ω be a Riemann matrix. For, referring to the corollary to Theorem 4 and considering the closed differential forms $\theta(z)$ where $\theta_j(z)$ are constants, we derive immediately that $-i\Xi$ is a positive definite Hermitian matrix. It is in this sense that we may refer to Theorem 4 as a generalization of the theorem of Frobenius. The converse implication would be of considerable interest, but nothing so strong is yet known.

The more general factors (11) have also been used previously. P. Myrberg has considered such factors in one complex variable in his studies of the analytic representation of automorphic functions [17]. In this connection, it will follow in the next section that whenever a factor (11) in one complex variable has a non-trivial character class, then all factors of automorphy are equivalent to rational multiples of that factor; hence the Myrberg factors are equivalent, insofar as representing automorphic functions abstractly, to the Poincaré factors which we shall consider next.

11. If D is a bounded subdomain of the complex affine space and Γ is a properly discontinuous group of analytc homeomorphisms of D onto itself, we may introduce the Poincaré factors of automorphy $\{J_T(z)\}$, where $J_T(z)$ denotes the complex Jacobian determinant of the mapping T. A differential form representing the character class of this factor in the sense of Theorem 5 can be constructed immediately. Considering T as a mapping on a real $2n$-dimensional Euclidean space, its real Jacobian determinant is $j_T(z) = |J_T(z)|^2$. If $ds^2 = \sum_{j,k} g_{jk}dx^j dx^k$ is a Γ-invariant real metric on D and $g(z) = \det(g_{jk}(z))$, then $[g(Tz)]^{\frac{1}{2}} = j_T(z)[g(z)]^{\frac{1}{2}}$; hence, as in Theorem 5, the desired differential form is just

$$(14) \qquad \phi(z) = \bar{\partial}\partial \log[g(z)]^{\frac{1}{2}} = \sum_{\alpha,\beta} R_{\alpha\bar{\beta}} dz^\alpha \wedge d\bar{z}^\beta,$$

where $R_{\alpha\bar{\beta}}$ is the complex Ricci curvature tensor. Consequently the necessary and sufficient condition that all factors of automorphy be equivalent to rational powers of the Poincaré factor is that all Γ-invariant differential forms $\bar{\partial}\partial g(z)$ with rational periods on the cycles $\Phi(v)$ of Section 1 be cohomologous to rational multiples of the complex Ricci curvature form (13). Although this condition is certainly not always fulfilled, for example when D

is the Cartesian product of domains D_1 and D_2, and Γ is the direct product of a group Γ_1 on D_1 and Γ_2 on D_2, nevertheless it is fulfilled trivially in one complex variable. Therefore in one complex variable all factors of automorphy are equivalent to rational powers of the Poincaré factors $J_T(z) = dT(z)/dz$.

The results in one complex variable can also be obtained directly from the known structural properties of the group. We may select a set of free generators $s_1, \cdots, s_p, t_1, \cdots, t_p, u_1, \cdots, u_q$ for H such that P is the normal subgroup generated by the words $u_1^{m_1}, \cdots, u_q^{m_q}$,

$$w = u_q \cdots u_1 s_p^{-1} t_p^{-1} s_p t_p \cdots s_1^{-1} t_1^{-1} s_1 t_1.$$

Then $P \cap [H, H]/[P, H]$ is the infinite cyclic group generated by the coset of $w^m (u_1^{-m_1})^{m/m_1} \cdots (u_q^{-m_q})^{m/m_q}$, where $m = \text{l.c.m.}(m_1, \cdots, m_q)$. The character class is therefore completely determined by its value on the generator of the group $P \cap [H, H]/[P, H]$, which may be called the characteristic number. If a factor of automorphy ν has a non-trivial character class, so that its characteristic number $M \neq 0$, then it follows from Theorem 2 that all factors of automorphy are equivalent to the factors $\nu_T(z)^{m/M}$ for integers m. In particular, since it follows from (13) that the Poincaré factors have non-trivial character class, all factors are equivalent to the factors $(dT(z)/dz)^{m/M}$. If no fixed points are present, it is well known that $M = 2p - 2$, where p is the genus of the Riemann surface D/Γ.

PRINCETON UNIVERSITY.

REFERENCES.

[1] P. Appell, "Sur les fonctions périodiques de deux variables," *Journal de Mathématiques Pures et Appliquées*, vol. 97 (1884), pp. 16-48.

[2] W. L. Baily, "On the quotient of an analytic manifold by a group of analytic homeomorphisms," *Proceedings of the National Academy of Sciences*, vol. 40 (1954), pp. 804-808.

[3] S. Bochner, "On compact complex manifolds," *Journal of the Indian Mathematical Society*, vol. 11 (1947), pp. 1-21.

[4] ———, and W. T. Martin, *Several Complex Variables*, Princeton, 1948.

[5] H. Cartan, "Variétés analytiques complexes et cohomologie," *Colloque sur les Fonctions de Plusieurs Variables*, Brussels, 1953.

[6] ———, *Séminaire École Normale Supérieure*, 1951-1952 (mimeographed notes,

reprinted by the Department of Mathematics, Massachusetts Institute of Technology).

[7] S. Eilenberg and S. MacLane, "Group extensions and homology," *Annals of Mathematics*, vol. 43 (1942), pp. 757-831.

[8] ———, "Cohomology theory in abstract groups," *Annals of Mathematics*, vol. 48 (1947), pp. 51-78.

[9] G. Frobenius, "Ueber die Grundlagen der Theorie der Jacobischen Funktionen," *Journal für die reine und angewandte Mathematik*, vol. 97 (1884), pp. 16-48.

[10] P. R. Garabedian and D. C. Spencer, "A complex tensor calculus for Kähler manifolds," *Acta Mathematica*, vol. 89 (1953), pp. 279-331.

[11] R. C. Gunning, "General factors of automorphy," *Proceedings of the National Academy of Sciences*, vol. 44 (1955), pp. 496-498.

[12] W. V. D. Hodge, *The Theory and Application of Harmonic Integrals*, Cambridge, 1952.

[13] H. Hopf, "Fundamentalgruppe und zweite Bettische Gruppe," *Commentarii Mathematici Helvetici*, vol. 14 (1941), pp. 257-309.

[14] K. Kodaira and D. C. Spencer, "Groups of complex line bundles over compact Kahler varieties," *Proceedings of the National Academy of Sciences*, vol. 39 (1953), pp. 868-872.

[15] B. O. Koopman and A. B. Brown, "On the covering of analytic loci by complexes," *Transactions of the American Mathematical Society*, vol. 34 (1932), pp. 231-251.

[16] S. Lefschetz and J. H. C. Whitehead, *Transactions of the American Mathematical Society*, vol. 34 (1933), pp. 510-528.

[17] P. Myrberg, "Sur les fonctions automorphes," *Annales Scientifiques de l'École Normale Supérieure*, vol. 68 (1951), pp. 383-424.

[18] G. deRham and K. Kodaira, *Harmonic Integrals* (mimeographed notes, Institute for Advanced Study, 1950).

[19] C. L. Siegel, *Analytic Functions of Several Complex Variables* (mimeographed notes, Institute for Advanced Study, 1949).

[20] A. Weil, "Théorèmes fondamentaux de la théorie des fonctions thêta," *Seminaire Bourbaki*, May, 1949.

RATIONAL EQUIVALENCE OF ARBITRARY CYCLES.*[1]

By Pierre Samuel.

We intend to give a definition and some properties of the notion of rational equivalence for cycles of arbitrary dimension on a non singular projective variety. That there exists such a theory of rational equivalence for arbitrary cycles was demonstrated by the work of F. Severi on the "series of equivalence." In this paper we will not attempt to make explicitly the connection with Severi's theory, but the influence of his work will be easy to detect. We will content ourselves to provide the working geometer with a certain number of tools. The definition of rational equivalence that we give is analogous to the definition of algebraic equivalence given by A. Weil ([6]), and the principal results we prove are also valid for algebraic equivalence, with trivial modifications in the proofs; it may even be observed that the only result in which we use non-elementary methods (i. e. the specialization theorem, in the proof of which we use both the degeneration principle and a property of the divisors of the second kind) has, in the case of algebraic equivalence, an analogue which is trivial. It goes without saying that many proofs have been inspired by those in A. Weil's paper ([6]). I am very grateful to J. I. Igusa, G. Washnitzer and O. Zariski for their valuable encouragement and their invaluable advice during the preparation of this paper.

1. Preliminary results. We use the terminology and notations of A. Weil ([5]), sometimes modified according to a recent book of ours ([4]). As we shall work only with *non singular* varieties, it will be sufficient, in order to prove that the intersection product $X \cdot Y$ of two cycles on a variety V is defined, to check that all the components of $\mathrm{Supp}(X) \cap \mathrm{Supp}(Y)$ [2] have the right dimension. We first recall some well known facts, which will be useful in the sequel:

(a) In order to apply the associativity formula to the intersection of, let us say, three cycles X, Y, Z on a variety V, it is sufficient to show that

* Received September 14, 1955.

[1] This work was supported by a research project at Harvard University, sponsored by the Office of Ordnance Research, United States Army, under contract DA-19-020-ORD-3100.

[2] By $\mathrm{Supp}(X)$ we mean the reunion of the components of the cycle X.

the intersection product obtained by "associating" X, Y, Z in a given order are defined.

(b) For the projection formula $Y \cdot \mathrm{pr}_U X = \mathrm{pr}_U((Y \times V) \cdot X)$ on the product of two non singular projective varieties to be valid, it is sufficient that its right hand side (i. e. $(Y \times V) \cdot X$) be defined.

Given a homogeneous cycle X on a variety V, it is convenient to consider its *codimension*, i. e. the integer $\dim(V) - \dim(X)$; we denote it by $\mathrm{cod}_V(X)$, or $\mathrm{cod}(X)$ when no confusion may arise. Then, if the intersection product $X \cdot Y$ of two cycles X, Y on V is defined, we have

(c) $\mathrm{cod}_V(X \cdot Y) = \mathrm{cod}_V(X) + \mathrm{cod}_V(Y)$.

If W is a subvariety of V such that $X \cdot W$ is defined, we have

(d) $\mathrm{cod}_W(X \cdot W) = \mathrm{cod}_V(X)$.

As in [4], Chap. I, § 10, No. 3, we do not restrict the notion of *rational mapping* of V into W to those mappings F such that $F(V) = W$. We identify the rational mapping F with its graph in $V \times W$; the fact that F is a rational mapping means that F is a variety, that $\mathrm{pr}_V(F) = V$, and that the projection index of F in V is equal to 1. We recall that, for a rational mapping F of V into W to be *regular* (i. e. regular at every point of V), it is necessary and sufficient that the birational correspondence between F and V defined by pr_V be *biregular* ([4]). It is easily seen that, if F is a rational mapping of V into W and if Y is a cycle on W such that $F^{-1}(Y) = \mathrm{pr}_V((V \times Y) \cdot F)$ is defined, then we have

(e) $\mathrm{cod}_V(F^{-1}(Y)) = \mathrm{cod}_W(Y)$

(cf. [4], Chap. II, § 6, No. 9, e)). It can also be proved, by using the same method as in Chap. II, § 6, No. 9, f), that, if F is a *regular* mapping of V into W, and if Y and Y' are two cycles on W such that $F^{-1}(Y)$, $F^{-1}(Y')$, $Y \cdot Y'$ and $F^{-1}(Y \cdot Y')$ are defined, then $F^{-1}(Y) \cdot F^{-1}(Y')$ is defined, and we have

(f) $F^{-1}(Y \cdot Y') = F^{-1}(Y) \cdot F^{-1}(Y')$.

It may be observed by using (a) that, at least in the case in which V and W are non singular, it is sufficient to assume that $Y \cdot Y'$ and $F^{-1}(Y \cdot Y')$ are defined, i. e. that $(V \times Y) \cdot (V \times Y') \cdot F$ is defined.

LEMMA 1. *Let S, R and V be three non singular projective varieties, Z a cycle in $V \times R$ and F a regular mapping of S into R. Then the mapping F' of $V \times S$ into $V \times R$ defined by $F'(v \times s) = v \times F(s)$ is regular. If a*

is a point of S such that $Z(F(a)) = \mathrm{pr}_V((V \times F(a)) \cdot Z)$ is defined, then the cycles $Z' = F'^{-1}(Z)$ on $V \times S$ and $Z'(a)$ in V are defined, and we have $Z'(a) = Z(F(a))$.

Proof. The regularity of F' is straightforward. As we shall work in the quadruple product $V \times S \times V \times R$, we shall, in order to avoid confusions, denote its third factor by V': $V \times S \times V' \times R$. If we denote by V^D the diagonal of the product $V \times V'$, the graph of F' is $F \times V^D (F \subset R \times S$, $V^D \subset V \times V')$. Our hypothesis that $Z(F(a)) = \mathrm{pr}_{V'}((V' \times F(a)) \cdot Z)$ is defined means that $(V' \times F(a)) \cdot Z$ is defined (intersection in $V' \times R$). We first prove that the intersection cycle

$$T = (V \times S \times Z) \cdot (V \times a \times V' \times R) \cdot (F \times V \times V') \cdot (S \times R \times V^D)$$

is defined. Since F is regular, $(a \times R) \cdot F$ is defined and equal to $a \times F(a)$; whence the intersection product of the two middle factors of T is defined, and is equal to $V \times a \times V' \times F(a)$. We are thus reduced to proving that

$$T = (V \times S \times Z) \cdot (V \times a \times V' \times F(a)) \cdot (S \times R \times V^D)$$

is defined. Since $(V' \times F(a)) \cdot Z$ is defined and equal to $Z(F(a)) \times F(a)$, the intersection product of the two first factors in T is defined and equal to $V \times a \times Z(F(a)) \times F(a)$. We are now reduced to prove that

$$T = (V \times a \times Z(F(a)) \times F(a)) \cdot (S \times R \times V^D)$$

is defined. Since this cycle, in $S \times R \times V \times V'$, is

$$(a \times F(a)) \times (V^D \cdot (V \times Z(F(a)))),$$

it is obviously defined, and its projection on V is $Z(F(a))$.

By the remark (a) above, all the partial intersection cycles in the formula for T are defined. In particular the intersection product of the first, third and fourth factors is defined; since it is equal to $(V \times S \times Z) \cdot (F \times V^D) = (V \times S \times Z) \cdot F'$, this shows that $Z' = F'^{-1}(Z) = \mathrm{pr}_{V \times S}((V \times S \times Z) \cdot F')$ is defined. On the other hand we have, by the projection formula

$$\mathrm{pr}_{V \times S}(T) = \mathrm{pr}_{V \times S}(((V \times S \times Z) \cdot F') \cdot (V \times a \times V' \times R)) = (V \times a) \cdot Z'.$$

This proves that $(V \times a) \cdot Z'$ is defined, whence also $Z'(a) = \mathrm{pr}_V((V \times a) \cdot Z')$. We therefore have $Z'(a) = \mathrm{pr}_V(\mathrm{pr}_{V \times S}(T))$, whence $Z'(a) = \mathrm{pr}_V(T) = Z(F(a))$.

LEMMA 2. *Let Z^z be a subvariety of the product $V^v \times W^w$ of two non singular varieties V, W, and let A^a be a subvariety of V. Denote by V_j the set of all points P of V such that $\dim(Z \cap (P \times W)) \geqq j$. Then V_j is a*

closed subset of V. *If* A *intersects properly all the components of all the closed sets* V_j, *then the intersection product* $Z \cdot (A \times W)$ *is defined* (*whence also* $Z(A) = \mathrm{pr}_W(Z \cdot (A \times W))$).

Proof. For every integer j such that $V_j \neq V_{j+1}$ we denote by A_j the point set $(V_j - V_{j+1}) \cap A$. The fact that V_j is a closed set follows from [4], Chap. I, § 8, No. 2, c). Then A_j is a finite union of open subvarieties of V, and, by hypothesis, each of them has a dimension $\leqq a - v + \dim(V_j)$. Since we have $\dim(Z(P)) = j$ for every P in $V_j - V_{j+1}$, the principle of counting constants ([4], Chap. I, § 10, No. 2) shows that the dimension of $Z \cap (A_j \times W)$ is $\dim(A_j) + j$ whence at most $a - v + \dim(V_j) + j$. We see in the same way that $\dim(Z \cap (V_j \times W)) = \dim(W_j) + j$, whence $\dim(W_j) + j \leqq \dim(Z) = z$ since $Z \cap (V_j \times W) \subset Z$. It follows that we have $\dim(Z \cap (A_j \times W)) \leqq a - v + z$. Since A is the union of the sets A_j (which are in finite number), $Z \cap (A \times W)$ is the union of the sets $Z \cap (A_j \times W)$, whence we have the inequality $\dim(Z \cap (A \times W)) \leqq a - v + z = z + (a + w) - (v + w)$. Since $z + (a + w) - (v + w)$ is the proper dimension for $\dim(Z \cap (A \times W))$, this proves our assertion as $V \times W$ is non singular.

LEMMA 3. *Let* V^n *be a non singular projective variety imbedded in* P_q, A^a *a subvariety of* V, (B_j) *a finite family of subvarieties of* V. *Then, for almost every linear variety* L^{q-n-1} *the projecting cone* C *of* A *with vertex* L *is such that, if we write* $C \cdot V = A + R$, *then the residual intersection* R *properly intersects all the varieties* B_j.

Proof. If $\dim(B_j) + a \geqq n$, let B_j' be a generic plane section of B_j of dimension $n - a - 1$; then R intersects properly B_j if and only if $R \cap B_j$ is empty. In other words, replacing B_j by B_j', we may assume that we have $\dim(B_j) + a < n$, and have to prove that L may be chosen in such a way that $R \cap B_j$ is empty for every j. Let W_j be the reunion of all the straight lines joining a point of A and a point of B_j (and of all their specializations; if (a) and $(b^{(j)})$ are affine generic points of A and B_j over a common field of definition k, and if t is a trancendental element over $k(a, b)$, then W_j is the locus of $(ta + (1 - t)b^{(j)})$ over k). The dimension of W_j is $\leqq a + \dim(B_j) + 1 \leqq n$. Thus almost all linear varieties L^{q-n-1} have an empty intersection with all the W_j's. For such a linear variety L, and for every point (a) of A, the linear variety L'^{q-n} containing L and (a) has at most (a) as common point with B_j. More precisely, if (b) is a point of $B_j \cap R$, then the linear variety L'^{q-n} determined by (b) and L intersects A at a point (a) which must coincide with (b), otherwise the line $(a)(b)$ would

meet L, in contradiction with the choice of L. Now, since the point $(a)=(b)$ is common to A and R, since $C\cdot V=A+R$, and since $(a)=(b)$ is a simple point of V, the tangent linear variety of V at $(a)=(b)$ must contain the projecting linear variety L'; then, since L' is tangent to V at a point common to A and B_j, it must contain a specialization of the line joining a generic point of A to a generic point of B_j, i.e., a line lying on the variety W_j; this again contradicts the choice of L. Therefore $B_j \cap R$ is empty.

2. Definition and characterizations of rational equivalence.

DEFINITION. *Let V^n be a non singular projective variety. A cycle X^r on V^n is said to be rationally equivalent to 0 if there exist a non singular unirational variety R^m, a cycle Z^{m+r} on $V\times R$ and two points a and b of R^m such that $Z(a)=\mathrm{pr}_V(Z\cdot(V\times a))$ and $Z(b)$ are defined and that $X=Z(a)-Z(b)$.*

We recall that a variety R is said to be *unirational* if its absolute function field is a subfield of a purely transcendental extension of the universal domain. We denote by $\mathfrak{R}_r(V)$ the set of all r-cycles on V which are rationally equivalent to 0.

THEOREM 1. *The set of cycles $\mathfrak{R}_r(V)$ is a group under addition.*

Proof. Let X and X' be two elements of $\mathfrak{R}_r(V)$. We write

$$X=Z(a)-Z(b), \qquad X'=Z'(a')-Z'(b')$$

where a, b (resp. a', b') are points of a non singular unirational variety R (resp. R'), and where Z (resp. Z') is a cycle on $V\times R$ (resp. $V\times R'$). We consider, on $V\times R\times R'$ the cycle $U=Z\times R'-R\times Z'$. Since

$$\begin{aligned} U(a\times a') &= \mathrm{pr}_V(U\cdot(V\times a\times a')) \\ &= \mathrm{pr}_V((Z\cdot(V\times a))\times a'-(Z'\cdot(V\times a'))\times a)=Z(a)-Z'(a'), \end{aligned}$$

and since, similarly, $U(b\times b')=Z(b)-Z'(b')$, we have

$$X-X'=U(a\times a')-U(b\times b').$$

As $R\times R'$ is a unirational variety, this proves that $X-X'\in\mathfrak{R}_r(V)$.

THEOREM 2. *If an r-cycle X on a non singular projective variety V^n is rationally equivalent to 0 there exist two points a, b of the projective line P_1 and a positive cycle T on $V\times P_1$ such that $T(a)$ and $T(b)$ are defined and that $X=T(a)-T(b)$.*

Proof. We write $X = Z(a) - Z(b)$, Z, a, b being as in the definition. We first reduce ourselves to the case in which Z is positive. We write $Z = Z^+ - Z^-$, where Z^+ and Z^- are the positive and the negative part of Z (cf. [4], Chap. I, § 9, No. 2, c), and we consider, on $R \times V \times R$, the cycle $U = Z^+ \times R + R \times Z^-$. By a simple computation as in Theorem 1, we see that

$$U(a \times b) = \mathrm{pr}_V(U \cdot (a \times V \times b)) = Z^+(a) + Z^-(b)$$

(these cycles being defined since $Z(a)$ and $Z(b)$ are defined). Similarly we have $U(b \times a) = Z^+(b) + Z^-(a)$, whence $X = U(a \times b) - U(b \times a)$. Since $R \times R$ is a unirational variety, we have achieved the reduction to the case of a positive cycle U.

If we now show that any two points of a unirational variety may be connected by a rational curve, the proof of Theorem 2 will be complete. In fact we shall have a rational mapping F of P_1 into $R \times R$ and two points c and d of P_1 such that $F(c) = a \times b$ and $F(d) = b \times a$. Since P_1 is a non singular curve, the local rings of all its points are valuation rings, and this proves that F is regular. Our conclusion then follows from Lemma 1, Section 1. We are thus reduced to proving the following lemma:

LEMMA 4. *Given two points a, b of a unirational variety V, there exists a rational curve lying on V and joining a and b.*

Proof of the lemma. The following terminology will be convenient. We say that a variety W *dominates* a variety U at a point u of U if there exists a rational mapping F of W onto U and a point w of W such that F is regular at w and that $F(w) = u$; we say that W dominates U if it dominates U at every point of U. According to Lüroth's theorem we may replace V by any rational variety which dominates V at a and b. Since there exists a rational mapping H of a projective space P_q onto V, we first replace V by the graph H, and we may thus assume the existence of a birational and regular mapping F of V onto P_q. As a second step we show the existence of a birational correspondence T between P_q and another projective space P_q' such that P_q' dominates V at a, and that P_q' corresponds biregularly to P_q at $F(b)$; let a' be a point of P_q' having a as regular image on V. If we apply the same result to P_q', the join V' of V and P_q', a' and $(b, T^{-1}(F(b)))$ instead of P_q, V, $F(b)$ and a, we obtain a projective space P_q'' which dominates V at a and b. Since any two points of a projective space may be joined by a straight line, this will prove the lemma.

Therefore we need only to show the existence of a birational correspondence T between P_q and P_q' such that P_q' dominates V at a and that P_q' corresponds biregularly to P_q at $F(b)$. We are immediately reduced to the affine case, where $a, b, F(a), F(b)$ are at finite distance. Let $x_1, \cdots, x_q$ be the coordinates in the affine space A_q, and $(x_1, \cdots, x_q, z_1, \cdots, z_q)$ a generic point of V over an algebraically closed field k; the quantities z_i are rational functions of $x_1, \cdots, x_q$ over k. We may assume that the points a and b are rational over k; and that $F(a)$ and a are the origins in A_q and A_{q+r}.

We shall apply the classical method of *successive quadration transformations* (analogous results are proved in the work of O. Zariski, by which our proof is directly inspired). We first notice the existence of a discrete valuation of rank 1 of $k(x_1, \cdots, x_q)$ which is zero-dimensional (i.e., since k is algebraically closed, the corresponding place ϕ takes its values in k), and which admits a as center on V. The existence of such an "analytical arc" is well known (see O. Zariski, "Foundations of a general theory of birational correspondences," *Transactions of the American Mathematical Society*, vol. 53 (1943), pp. 490-542; this statement is proved on pp. 501-502 as case (a) in Theorem 5; the method of proof shows that the constructed valuation may be assumed to be discrete). The numbers $v(x_j)$, $v(z_i)$ are then > 0. After a suitable linear change of coordinates we may assume that $F(b)$ does not lie on $X_1 = 0$, that $v(x_1)$ is the smallest of the numbers $v(x_j)$, and (replacing x_j by $x_j - \phi(x_j/x_1)x_1$) that $v(x_j) > v(x_1)$ for $j \geqq 2$.

The quadratic transformation $x_1 = x_1'$, $x_j = x_1'x_j'$ for $j \geqq 2$ is then biregular at $F(b)$. For any polynomial $D(x)$ we have $D(x) = (x_1')^d D'(x')$, where d is the order of $D(x)$ (i.e. the degree of its lowest degree form), and where $D'(x')$ is a polynomial uniquely determined by $D(x)$. If d is $\neq 0$, we have $v(D'(x')) < v(D(x))$ and $v(x_1') \leqq v(D(x))$. We say that a finite family of polynomials is adequate if each rational function z_i is a power product of these polynomials. Let $(A_1(x), \cdots, A_s(x))$ be such a family. Then z_i, considered as a rational function of the variables (x'), is a power product of the polynomials x_1', $A_u'(x')$, and the family $(x_1', A_1'(x'), \cdots, A_s'(x'))$ is adequate (for the variables (x')). If one at least of the orders $v(A_u(x))$ is $> v(x_1)$, we have the inequality

$$\text{(I)} \qquad \max(v(x_1'), v(A_u'(x'))) < \max(v(A_u(x))).$$

Since v is a discrete valuation of rank 1, we cannot repeat this procedure of quadratic transformations an infinite number of times and always get an inequality of the type (I) between the orders (for v) of the elements of two successive adequate families. Therefore we eventually get a system $(y_1, \cdots, y_q)$

of independent variables (with $v(y_j) > 0$ for every j), and an adequate family $(B_u(y))$ such that either $v(B_u(y)) = 0$ (in which case $B_u(y)$ is a polynomial with non-zero constant term), or $v(B_u(y)) = v(y_1)$. By a last quadratic transformation $y_1 = y_1'$, $y_j = y_1 y_j'$ for $j \geqq 2$ (we assume as above that $v(y_j) > v(y_1)$), we get, for the indices u such that $v(B_u(y)) = v(y_1)$, $B_u(y) = y_1' B_u'(y')$, and $B_u(y) = B_u'(y')$ for the others; at any rate all the polynomials $B_u'(y')$ have a constant term $\neq 0$. Hence we can write $z_i = (y_1')^{m(i)} R_i(y')$, where $R_i(y')$ is a rational function belonging to the local ring $\mathfrak{o}$ of the origin in the affine space with coordinates (y'). On the other hand the elements x_j are polynomials in the variables (y') by construction. Since the numbers $v(z_i)$, $v(y_1')$ are > 0, we have $m(i) > 0$ for every i. Thus all the elements z_i, x_j belong to the local ring $\mathfrak{o}$.

THEOREM 3. *Let X^r be an r-cycle on a non singular projective variety V^n. For X to be rationally equivalent to* 0 *on V it is necessary and sufficient that there exist a Chow variety W of positive r-cycles on V, a rational curve C on W and two points a, b of C such that $X = c(a) - c(b)$ ($c(x)$ denoting the cycle corresponding to the Chow point x).*

Proof. We first prove the necessity of our condition. Since X is rationally equivalent to 0, we can write $X = Z(c) - Z(d)$, where Z is a positive cycle on $V \times P_1$, and where c, d are points of P_1 (Theorem 2). Let t be a generic point of P_1 over an algebraically closed field of definition k of c, d, V and Z. Since $Z(a)$ and $Z(b)$ are specializations of $Z(t)$ over k ([4], Chap. II, § 6, No. 8, a), the Chow points ([4], Chap. I, § 9, No. 6) belong to a common Chow variety W of positive cycles (and W is defined over k; cf. [4], Chap. I, § 9, No. 6). Since the Chow point x of $Z(t)$ is rational over $k(Z(t)) \subset k(t)$, the locus C of x over k is a rational curve by Lüroth's theorem. Taking for a and b the Chow points of the cycles $Z(c)$ and $Z(d)$, we immediately see that a and b lie on C. Thus the necessity is proved.

We now prove the sufficiency. *Let* $X = c(a) - c(b)$, where a and b denote two points of a rational curve C lying on a Chow variety W of positive cycles on V. Let k be an algebraically closed field of definition of V, W, C, a, b, and let x be a generic point of C over k. The cycle $c(x)$ is rational over some purely inseparable extension K of $k(x)$ ([4], Chap. I, § 9, No. 4, g). As k is perfect and as $k(x)$ is a simple transcendental extension of k, say $k(t')$ (the system (t') being reduced to one element), there exists an exponent e such that $K = k(t'^{p^{-e}})$. Thus K is a simple transcendental extension $k(t)$ ($t = t'^{p^{-e}}$). Since the cycle $c(x)$ is rational over $k(t)$, we may consider its

locus Z in $V \times P_1$ over k ([4], Chap. II, § 6, No. 8, b): $\mathrm{pr}_V((V \times t) \cdot Z) = c(x)$. If F denotes the regular mapping of $V \times P_1$ onto $V \times C$ such that $F(v \times t) = x \times c(x)$, the point set Z is the inverse image $F^{-1}(T)$ of the incidence correspondence T attached to the system C of cycles ([4], Chap. I, § 10, No. 4, b) (as to cycles we have $F^{-1}(T) = p^f \cdot Z$). We denote by c and d two points of P_1 which are mapped into a and b by the restriction of F to P_1. Since $T(a)$ and $T(b)$ (considered as point sets) have the right dimension and are equal, as cycles, to $p^f \cdot c(a)$ and $p^f \cdot c(b)$, Lemma 1, Section 1 shows that the cycles $p^f \cdot Z(c)$ and $p^f \cdot Z(d)$ are defined and equal to $p^f \cdot c(a)$ and $p^f \cdot c(b)$, whence $X = Z(c) - Z(d)$. Therefore the sufficiency is proved since P_1 is a unirational variety.

We notice that, if a *divisor* X^{n-1} on V^n is rationally equivalent to 0, then a theorem proved in [5] (Chap. VIII, No. 2, Th. 5) shows that it is linearly equivalent to 0. The converse is obvious.

We say that two r-cycles X, X' on a non singular projective variety V are *rationally equivalent* on V if their difference $X - X'$ is rationally equivalent to 0 (i.e. if $X - X' \varepsilon \mathfrak{R}_r(V)$). Since $\mathfrak{R}_r(V)$ is a group (Theorem 1), this is actually an equivalence relation. Since, in the case of divisors, it coincides with linear equivalence, we may denote it by the same symbol, and write $X \sim X'$.

3. Some properties of rational equivalence.

THEOREM 4 (*"specialization theorem"*). *Let X^r be an r-cycle on a projective non singular variety V, which is rationally equivalent to 0 on V. If X' is a specialization of X over some field of definition k of V, then X' is rationally equivalent to 0 on V.*

Proof. Since every specialization over k is the same thing as a specialization over the algebraic closure of k followed by a k-automorphism, and since everything algebraic is preserved by k-automorphisms, we may assume that k is algebraically closed. By Theorem 3 there exists a Chow variety W of r-cycles on V, a rational curve C on W, and two points a and b of C such that $X = c(a) - c(b)$. We extend the specialization $X \rightarrow X'$ to a k-specialization $(X, C, a, b) \rightarrow (X', C', a', b')$. Since W is defined over k and is complete, C' is a 1-cycle on W, and a', b' are two points of C'. If we show that C' is connected and that all its components are rational curves, the proof of Theorem 4 will be complete: in fact we have $X' = c(a') - c(b')$, and there exist rational curves $C_1', \cdots, C_n'$ on W and connecting points $d_1, \cdots, d_{n-1}$ such that $a', d_1 \varepsilon C_1', d_1, d_2 \varepsilon C_2', \cdots, d_{n-1}, b' \varepsilon C_n'$; then the cycles $c(a') - c(d_1)$,

$c(d_1)-c(d_2),\cdots,c(d_{n-1})-c(d')$ are rationally equivalent to 0 by Theorem 3, whence also their sum X' by Theorem 1. We are thus reduced to proving:

LEMMA 5. *If a 1-cycle C' (in some projective space P_n) is a specialization over k of a rational curve C, then C' is connected, and all its components are rational curves.*

That C' is connected follows from the degeneration principle ([8]). Replacing C by a generic specialization $\bar{C}$ over k, and denoting by $(\bar{c})$ and (c') the Chow points of $\bar{C}$ and C', Prop. 7 in App. II of [5] shows that there exists an algebraically closed field K containing k, a set of quantities (x), a non singular curve U defined over K and admitting $(x,\bar{c})$ as generic point over K, and a point (x',c') of U: in fact we may suppose that (c') is a non generic specialization of (c) (otherwise our assertion is trivial), and the facts that we may take K algebraically closed and U non singular (i. e. normal over K) follow from the proof of the above quoted Prop. 7. Since $\bar{C}$ is irreducible, all the coefficients of the components of $\bar{C}$ are prime to the characteristic (there is only one of them, which is equal to 1), and a result of Chow (cf. [7], App., Prop. 6) shows that $\bar{C}$ is a rational curve over $k(\bar{c})$; whence $K(x,\bar{c})$ is a field of definition of the curve $\bar{C}$. Let (y) be a generic point of $\bar{C}$ over $K(x,\bar{c})$. Let S $(\subset P_q\times U)$ be the surface locus of $(y,\bar{c},x)$ over K. Its horizontal section $(P_q\times(c',x'))\cap S$ is equal to C', at least as a point set.

Since $\bar{C}$ is a rational curve, $K(x,\bar{c})(y)$ is a function field of genus 0 over $K(x,\bar{c})$. Since K is algebraically closed, and since $\dim_K(K(x,\bar{c}))=1$, $K(x,\bar{c})$ is quasi algebraically closed by a theorem of Tsen. Thus $K(x,\bar{c})(y)$ is a simple transcendental extension $K(x,\bar{c})(t)$ of $K(x,\bar{c})$, since a field of genus 0 admits a positive rational divisor of degree 2, and since a conic over a quasi-algebraically closed field admits a rational point. In other words the surface S is birationally equivalent to the product $P_1\times U$ (this is a well known result of Max Noether): more precisely there exists a birational correspondence T between S and $P_1\times U$ such that $T(x,\bar{c},y)=(x,\bar{c},t)$.

We have to prove that every component D of $(P_q\times(c',x'))\cap S$ is a rational curve. It is clear that every subvariety of $P_1\times U$ which corresponds to D under T lies on the straight line $P_1\times(x',c')$. We consider a prime divisor v of the function field of S (over K) having D as a center on S. If the center of v on $P_1\times U$ is the entire line $P_1\times(x',c')$, then the residue field R_v of the valuation v is the function field of this line, whence is a purely transcendental extension of K (the point (x',c') is rational over K since K is algebraically closed, and since it is non-generic specialization of $(x,\bar{c})$). Otherwise the center of v on $P_1\times U$ is a simple point of $P_1\times U$ (then the

prime divisor v is of the second kind on $P_1 \times U$), and a result about divisors of the second kind on surfaces ([1]) shows that the residue field R_v is also a purely transcendental extension of K in this case. Since R_v is the function field of some curve D^0 corresponding to D in a derived normal model S^0 of S, this proves that D^0 is a rational curve. Since D is a "projection" of D^0, D is also a rational curve by Lüroth's theorem. Lemma 5 and Theorem 4 are thereby proved.

THEOREM 5. *Let X and Y be two cycles of dimension r and s on a non singular projective variety V^n. If $X \sim 0$ and if $X \cdot Y$ is defined, then $X \cdot Y \sim 0$.*

Proof. We write $X = Z(a) - Z(b)$, where a and b are simple points of a rational variety R, and where Z is a cycle on $V \times R$ such that $Z(a)$ and $Z(b)$ are defined. We first study a particular case:

LEMMA 6. *If $Z(a) \cdot Y$ and $Z(b) \cdot Y$ are defined, then $X \cdot Y$ is rationally equivalent to 0 on V. If, furthermore, Y is a non singular subvariety of V, then $X \cdot Y$ is rationally equivalent to 0 on Y.*

In fact, if $Z(a) \cdot Y = Y \cdot \mathrm{pr}_V((V \times a) \cdot Z)$ is defined, then the set $(V \times a) \cap \mathrm{Supp}(Z) \cap \mathrm{Supp}(Y \times R)$, which is in biregular correspondence with $\mathrm{Supp}(Y) \cap \mathrm{Supp}(Z(a))$ has the correct dimension, whence (Section 1, (a)) $(V \times a) \cdot Z \cdot (Y \times R)$ and $Z' = Z \cdot (Y \times R)$ are defined. By the projection formula we have $Z(a) \cdot Y = \mathrm{pr}_V((V \times a) \cdot Z \cdot (Y \times R)) = Z'(a)$, whence $X = Z'(a) - Z'(b)$, and this proves our first assertion. For the second one we consider Z' as a cycle on $Y \times R$, and we still have $Z'(a) = \mathrm{pr}_Y((Y \times a) \cdot Z')$. This proves the lemma.

For completing the proof of Theorem 5, we take a field of definition k of all the components of V, R, a, b, X, Y and Z, and we consider Y as a specialization over k of a cycle $\bar{Y}$ such that $\bar{Y} \cdot Z(a)$ and $\bar{Y} \cdot Z(b)$ are defined: for example we take a projecting cone C of Y whose vertex is a linear variety L^{q-n-1} generic over k (q: dimension of a projective space in which V^n is imbedded), a generic projective transform $\bar{C}$ of C (over $k(C)$), and set $\bar{Y} = \bar{C} \cdot V + D$, where D is the residual intersection $C \cdot V - Y$ (Section 1, Lemma 3). Then $\bar{Y} \cdot X$ is rationally equivalent to 0 by Lemma 6. Since we are in a case where the specialization theorem ([4], Chap. II, § 6, No. 7) may be applied separately to positive and negative parts of the cycles we consider, $Y \cdot X$ is a specialization of $\bar{Y} \cdot X$, whence is ~ 0 on V by Theorem 4.

We shall prove later that the second assertion of Lemma 5, is still true even if $Z(a) \cdot Y$ and $Z(b) \cdot Y$ are not defined ($X \cdot Y$ being, of course, defined).

COROLLARY 1. *The group* $\mathfrak{R}_r(V)$ *contains the intersection cycles* $D_1 \cdot D_2 \cdot \cdots \cdot D_{n-r}$ *which are defined and such that one of the divisors* D_i *is* ~ 0.

COROLLARY 2. *For an r-cycle X in projective space* P_n *to be rationally equivalent to* 0 *in* P_n, *it is necessary and sufficient that it be of degree* 0.

In fact we use a theorem of Severi ([4], Chap. II, § 6, No. 4) and we write X as an intersection $X = D_1 \cdot \cdots \cdot D_{n-r}$ of divisors. If X has degree 0, then one of the D's must have degree 0 by Bezout's theorem; whence one of the D_i's is ~ 0, and also X by Theorem 5. For the converse we write $X = Z(a) - Z(b)$, and notice that $Z(a)$ and $Z(b)$ have the same degree.

COROLLARY 3. *A cycle X on V which is a complete intersection* (*i. e.* $X = V \cdot Y$, *Y being a cycle in the ambiant projective space* P_q) *is rationally equivalent to* 0 *on V if and only if it is of degree* 0 (*as a cycle in* P_q).

In fact the part "only if" is clear. Conversely, if X has degree 0, then Y has also degree 0 by Bezout's theorem. By Corollary 2 we may write $Y = D_1 \cdot \cdots \cdot D_s$ where the D_i's are divisors in P_q such that D_1 is of degree 0. Then the $X' = D_1^+ \cdot D_2 \cdot \cdots \cdot D_s \cdot V$, $X'' = D_1^- \cdot D_2 \cdot \cdots \cdot D_s \cdot V$ are defined, and since $D_1 = D_1^+ - D_1^-$ is a representation of D_1 under the form $Z(a) - Z(b)$, the relation $X = X' - X''$ is a representation of X under the form $Z'(a) - Z'(b)$. The conclusion follows by Lemma 5.

THEOREM 6. *Let V, W be two non singular projective varieties, and X a cycle on V which is* ~ 0 *on V. Then* $X \times W \sim 0$ *on* $V \times W$.

Proof. We write $X = Z(a) - Z(b)$ (a, b simple points of unirational variety R, Z cycle in $V \times R$). We set $Z' = Z \times W$ (in $V \times W \times R$). Then, by the formula on intersections on product varieties ([4], Chap. II, § 6, No. 5, f), $Z'(a)$ is defined and equal to $Z(a) \times W$. Thus $X \times W = Z'(a) - Z'(b)$.

COROLLARY. *If X is* ~ 0 *on V and if Y is any cycle on W, then* $X \times Y \sim 0$ *on W*.

Analogous proof by using $Z'' = Z \times Y$. Or notice that

$$X \times Y = (X \times W) \cdot (V \times Y),$$

and use Theorems 6 and 5.

THEOREM 7. *Let V, W be two non singular projective varieties, and let X be a cycle on* $V \times W$. *If* $X \sim 0$ *on* $V \times W$, *then* $\mathrm{pr}_V(X) \sim 0$ *on V*.

Proof. We write $X = Z(a) - Z(b)$ (a, b simple points of unirational variety R, Z cycle on $V \times W \times R$). We have

$$\mathrm{pr}_V(Z(a)) = \mathrm{pr}_V(\mathrm{pr}_{V\times W}(V \times W \times a) \cdot Z)$$
$$= \mathrm{pr}_V(\mathrm{pr}_{V\times R}((V \times W \times a) \cdot Z) = \mathrm{pr}_V((V \times a) \cdot \mathrm{pr}_{V\times R}(Z))$$

(we apply the projection formula, as recalled in (b), § 1). Thus, if we set $Z' = \mathrm{pr}V \times R(Z)$, we have $\mathrm{pr}_V(X) = Z'(a) - Z'(b)$.

THEOREM 8. *Let V and W be two non singular projective varieties, F a rational mapping of V into V, and X a cycle on V which is ~ 0 on V. Then, if $F^{-1}(X)$ is defined, it is ~ 0 on W.*

Proof. We have $F^{-1}(X) = \mathrm{pr}_W((W \times X) \cdot F)$, where $(W \times X) \cdot F$ is defined. Since $X \sim 0$, we have $W \times X \sim 0$ by Theorem 6, whence $(W \times X) \cdot F \sim 0$ on $W \times V$ by Theorem 5. Therefore $\mathrm{pr}_W((W \times X) \cdot F) \sim 0$ on W by Theorem 7.

It may be observed that, if Z is any cycle on $W \times V$, and X a cycle on V which is ~ 0 on V, then $Z(X) = \mathrm{pr}_W((W \times X) \cdot Z)$ is ~ 0 on W if it is defined. The proof is the same as in Theorem 8.

COROLLARY. *Let W be a non singular subvariety of a non singular projective variety V. If X is a cycle on V which is ~ 0 on V, and if $W \cdot X$ is defined, then $W \cdot X$ is ~ 0 on W.*

In fact, if we denote by i the inclusion mapping of W into V, we have $W \cdot X = i^{-1}(X)$.

Remark. It is *not* true that, if Y is a cycle on W ($\subset V$) which is ~ 0 *on* V, then it is ~ 0 *on* W: take $V = P_3$, W a quadric, Y the difference of two straight lines of different systems on W. But the converse is true: if Y is a cycle on W ($\subset V$) which is ~ 0 *on* W, then Y is ~ 0 on the bigger variety V. This follows from Theorem 3 and from the fact that a Chow variety of positive r-cycles on W is a subvariety of a Chow variety of r-cycles on V; or one may apply the remark following Theorem 8 to $i(Y)$.

We may use Lemma 5 (about specializations of rational curves) to prove a more general result than Theorem 4:

THEOREM 9. *Let X^r be an r-cycle on a non singular projective V^n, and let (X', V') be a specialization of (X, V) over some field k, such that the cycle V' is a non singular variety. If $X \sim 0$ on V, then $X' \sim 0$ on V'.*

Proof. It is well know that the relation "$\operatorname{Supp}(A) \subset \operatorname{Supp}(B)$" between two cycles A, B may be expressed by a system of equations with absolutely algebraic coefficients in the Chow coordinates of A and B ([2]; or [4], Chap. I, § 9, No. 7, g where an analogous result is proved). Thus X' is a cycle on V'. We now use Theorem 3 and write $X = c(a) - c(b)$ where a, b are points of a Chow variety W of positive r-cycles on V, such that a and b can be connected by a rational curve C lying on W. We extend the given specialization to $(V, X, W, C, a, b) \to (V', X', W', C', a', b')$. By [4], Chap. I, § 9, No. 7, h every point of $\operatorname{Supp}(W')$ is a specialization of a point of W; thus every point of $\operatorname{Supp}(W')$ is the Chow point of some positive r-cycle on V'; whence also a', b', and every point of $\operatorname{Supp}(C')$. Since C' is a connected union of rational curves (Lemma 5), it follows, as in the proof of Theorem 4, that $c(a') - c(b')$ is ~ 0 on V'. Since the relation $X = c(a) - c(b)$ gives, by specialization, $X' = c(a') - c(b')$, we have $X' \sim 0$ on V'.

THEOREM 10. *Let V^n be a non singular variety in projective space P_q. If a cycle X of dimension r is ~ 0 on V, then there exists a function f on V and a cycle Y^{r+1} on V such that $X = (f) \cdot Y$.*

Proof. We first prove that X may be written as $X = \sum_i (f_i) \cdot Y_i$. We write $X = Z(a) - Z(b)$, with $Z \subset V \times P_1$, $a, b \in P_1$. By decomposition of the cycle Z in irreducible components, we are reduced to the case in which Z is irreducible. Then $\operatorname{pr}_V(Z)$ is an irreducible subvariety W of V, whose dimension is r or $r+1$. If $\dim(W) = r$, then $Z(a) = Z(b) = W$, $X = 0$, and there is nothing to prove. Otherwise $Z(a)$ and $Z(b)$ are divisors on W, and, since their Chow-points may be connected by a rational curve of "W-divisors $Z(t)$," they are linearly equivalent on W. Denoting by H a suitable divisor of degree 0 in P_q, we thus have

$$X = Z(a) - Z(b) = (H \cdot W)_{P_q} = ((H \cdot V)_{P_q} \cdot W)_V.$$

Since $(H \cdot V)_{P_q}$ is the divisor of a function f on V, the first part of the proof is complete.

We may now write $X = \sum_i (f_i) \cdot Y_i^{r+1}$ where the Y_i's are distinct subvarieties of V, whence $X = \sum_i H_i \cdot Y_i$, where the H_i's are divisors of degree 0 in P_q. Let $P_i(x) = 0$ be the equation of H_i, $P_i(x)$ being the quotient of two forms of like degree. We can find forms $F_j(x)$ of the same degree such that F_j is $\neq 0$ on Y_j but is 0 on Y_i for $i \neq j$. Then the divisor H with equation $F(x) = (\sum_j F_j(x))^{-1}(\sum_j F_j(x) P_j(x)) = 0$ has degree 0 and is such

that $H \cdot Y_j = H_j \cdot Y_j$ for every j: in fact F and F_j induce the same function on Y_j. If f_j denotes the function induced by F on V, we have

$$X = \sum_j (f) \cdot Y_j = (f) \cdot \sum_j Y_j.$$

We conclude this section by showing that our notion of rational equivalence coincides with the notion of "linear equivalence of cycles of arbitrary dimension" defined by A. Weil in [5], p. 959 (bottom). We use here the word "rational" instead of "linear" since it has been used by F. Severi in a long series of papers for denoting a notion which is also the same as our notion, whereas A. Weil devotes only two lines to the "linear" equivalence of cycles of arbitrary dimension.

First it is clear that our groups $\mathfrak{R}_r(V)$ satisfy conditions (A), (B), (C′), (D), (E) and (L) of Weil. These properties are respectively stated in Theorem 6, Cor. to Theorem 8, Theorem 7, Theorem 5, Theorem 7, except for (L) which is evident. Thus our groups $\mathfrak{R}_r(V)$ contain the corresponding groups of Weil.

Conversely, if a cycle X on V is rationally equivalent to 0 (in our sense), we have $X = Z(a) - Z(b) = \mathrm{pr}_V(Z \cdot (V \times ((a) - (b))))$, where a and b are two points of P_1 and Z a cycle on $V \times P_1$. In the sense of Weil the cycles $(a) - (b), V \times ((a) - (b)), Z \cdot (V \times ((a) - (b)))$ and X are "linearly" equivalent to 0" by (L), (A), (D) and (E) respectively. This proves the opposite inclusion.

4. The ring of rational equivalence classes.

THEOREM 11. *Let V be a non singular projective variety, and let α and β be two rational equivalence classes of cycles on V. There exist cycles A and B in α and β such that $A \cdot B$ is defined. And the rational equivalence class of $A \cdot B$ depends only on α and β.*

Proof. We choose A and B' arbitrarily in α and β. If $A \cdot B'$ is not defined, we consider a common field of definition k of V, A, B', a generic linear variety L^{q-n-1} (q: dimension of the ambiant projective space) over k, the projecting cone C of B' with vertex L, and a generic projective transform $\bar{C}$ of C over $k(C)$. We set $C \cdot V = B' + D$, and $\bar{C} \cdot V = E$. The cycle $A \cdot D$ is defined by Lemma 3, § 1, and $A \cdot E$ is evidently also defined. We now take $B = E - D$; then $A \cdot B$ is defined. Since the group of projective transformations is a rational variety, we have $C \sim \bar{C}$ on P_q, whence $C \cdot V \sim \bar{C} \cdot V$ on V (Cor. to Theorem 8). In other words we have $B' + D \sim E$ on V, i.e., $B' \sim B$. This proves our first assertion.

Let now A, A' be elements of α and B, B' be elements of β such that $A \cdot B$ and $A' \cdot B'$ are defined. The first part of the proof shows the existence of B'' in β such that $A \cdot B''$ and $A' \cdot B''$ are defined. Then Theorem 5 shows that $A \cdot B \sim A \cdot B''$, that $A \cdot B'' \sim A' \cdot B''$, and that $A' \cdot B'' \sim A' \cdot B'$. Whence $A \cdot B \sim A' \cdot B'$ by transitivity. Therefore the rational equivalence class of $A \cdot B$ does not depend on the choice of A and B in α and β.

It follows from Theorem 11 that we have defined, on the set $\mathfrak{E}(V)$ of rational equivalence classes of cycles on V, a *multiplication*. The commutativity and the associativity of intersection products show that this multiplication is commutative and associative. We have also, on $\mathfrak{E}(V)$, an addition, defined by the addition of cycles; this addition is distributive with respect to the multiplication. The set $\mathfrak{E}_r(V)$ of rational equivalence classes of cycles of *codimension* r is an additive subgroup of $\mathfrak{E}(V)$, $\mathfrak{E}(V)$ is the direct sum of the $\mathfrak{E}_r(V)$, and, by (c), § 1, the product of an element of $\mathfrak{E}_r(V)$ and of an element of $\mathfrak{E}_s(V)$ is in $\mathfrak{E}_{r+s}(V)$. In other words $\mathfrak{E}(V)$, graded by the codimension, becomes a *graded commutative ring*; we call it the *ring of rational equivalence classes* on V. It admits a unit element, the equivalence class of V itself.

It is easily seen that the rational equivalence classes of the cycles which are *algebraically equivalent* to 0 on V form an *ideal* $\mathfrak{A}$ in $\mathfrak{E}(V)$. The factor ring $\mathfrak{E}(V)/\mathfrak{A}$ is canonically isomorphic to the ring of algebraic equivalence classes on V (defined by a method similar to the one used here for rational equivalence).

Examples. 1) If V is the projective space P_n, then the graded ring $\mathfrak{E}(V)$ is isomorphic to $Z[X]/(X^{n+1})$, the coset of X corresponding to the class of the hyperplanes (cf. Cor. 2 to Theorem 5).

2) If V is a quadric in P_3, $\mathfrak{E}(V)$ is isomorphic to $Z[X, Y]/(X^2, Y^2)$, the cosets of X and Y corresponding to the classes of the straight lines on V.

Given two non singular projective varieties V and W, and a *correspondence* Z between V and W (i. e. a cycle on $V \times W$), we are going to define an *additive mapping* Z^* of $\mathfrak{E}(W)$ into $\mathfrak{E}(V)$. In fact, given a rational equivalence class α on W, there exists a cycle A in α such that

$$Z(A) = \mathrm{pr}_V((V \times A) \cdot Z)$$

is defined: in fact the proof of Theorem 11 shows that, given a finite number of subvarieties (B_j) of W, there exists A in α such that $A \cdot B_j$ is defined for every j; thus our assertion follows from Lemma 2, § 1. On the other hand,

if A and A' are elements of α such that $Z(A)$ and $Z(A')$ are defined, then we have $Z(A) \sim Z(A')$ on V, by the remark following Theorem 8.

The mapping Z^* is *not* a ring homomorphism in general. However, if F is a *regular mapping* of V into W, we may use the formula $F^{-1}(A \cdot B) = F^{-1}(A) \cdot F^{-1}(B)$ recalled in § 1, (f). More precisely, given two rational equivalence classes α and β on V, we first choose A in α such that $F^{-1}(A)$ is defined (by Lemma 2, § 1 this amounts to requiring that A intersects properly a finite number of subvarieties W_j of W). We have now to choose B in β in such a way that $A \cdot B$, $F^{-1}(B)$ and $F^{-1}(A \cdot B)$ are defined; this means that B must intersect properly all the components of A, all the subvarieties W_j and all the components of $\mathrm{Supp}(A) \cap W_j$ for every j; this again is a requirement of proper intersection with a finite number of subvarieties of V, and this requirement may be fulfilled as in the proof of Theorem 11. Therefore the additive mapping F of $\mathfrak{E}(W)$ into $\mathfrak{E}(V)$ defined by F^{-1} satisfies $F^*(\alpha \cdot \beta) = F^*(\alpha) \cdot F^*(\beta)$ for every α, β. In other words F^* is a *homomorphism* of the ring $\mathfrak{E}(W)$ into $\mathfrak{E}(V)$, and this homomorphism preserves codimensions by § 1 (e), i. e. is *a homomorphism for the structures of graded rings of* $\mathfrak{E}(W)$ and $\mathfrak{E}(V)$.

If U, V, W are three non singular projective varieties, and $F: U \to V$ and $G: V \to W$ are *regular mappings*, then the composition $G \circ F$ is a regular mapping of U into W. It follows immediately from [4], Chap. II, § 6, No. 9, g that, when they are defined, the cycles $F^{-1}(G^{-1}(A))$ and $(G \circ F)^{-1}(A)$ (A: cycle on W) are equal. Consequently we have

$$F^*(G^*(\alpha)) = (G \circ F)^*(\alpha) \text{ for every } \alpha \text{ in } \mathfrak{E}(W).$$

In other words the homomorphisms F^* are *transitive*. We can also say that $\mathfrak{E}$ is a *contravariant functor* for the categories of non singular projective varieties and of regular mappings.

UNIVERSITE DE CLERMONT-FERRAND AND
HARVARD UNIVERSITY.

REFERENCES.

[1] S. Abhyankar, "On the valuations centered in a local domain," (to appear).

[2] W.-L. Chow and B. L. van der Waerden, "Uber zugeordnete Formen und algebraische systeme von algebraischen mannigfalligkeiten," *Mathematische Annalen*, vol. 113 (1937), pp. 692-704.

[3] T. Matsusaka, "The theorem of Bertini on linear systems in modular fields," *Memoirs College Science Kyoto University*, vol. 26 (1950), pp. 51-62.

[4] P. Samuel, "Méthodes d'Algèbre Abstraite en géométrie Algébrique," *Ergebnisse der Mathematik und ihrer Grenzgebiete*, Neue Folge, Heft 4 (1955), Berlin (Springer).

[5] A. Weil, *Foundations of Algebraic Geometry*, American Mathematical Society Colloquium Publications, 1946.

[6] ———, "Sur les critères d'équivalence en géométrie algébrique," *Mathematische Annalen*, vol. 128 (1954), pp. 95-127.

[7] ———, "On algebraic groups of transformations," *American Journal of Mathematics*, vol. 77 (1955), pp. 355-391.

[8] O. Zariski, "Holomorphic functions," *Memoirs of the American Mathematical Society*, No. 5 (1951).

[9] ———, "Pencils on an algebraic variety and a new proof of the theorem of Bertini," *Transactions of the American Mathematical Society*, vol. 50 (1941), pp. 48-70.

SOME BASIC THEOREMS ON ALGEBRAIC GROUPS.*

By Maxwell Rosenlicht.[1]

The subject of algebraic groups has had a rapid development in recent years. Leaving aside the late research by many people on the Albanese and Picard variety, it has received much substance and impetus from the work of Severi on commutative algebraic groups over the complex number field, that of Kolchin, Chevalley, and Borel on algebraic groups of matrices, and especially Weil's research on abelian varieties and algebraic transformation spaces. The main purpose of the present paper is to give a more or less systematic account of a large part of what is now known about general algebraic groups, which may be abelian varieties, algebraic groups of matrices, or actually of neither of these types.

The first two parts of our work are devoted largely to extending, by very similar methods, the work of Nakano and Weil on the construction of transformation spaces, homogeneous spaces, and factor groups. Our third part proves the expected homomorphism theorems, the fourth gives a useful result on the existence of cross sections, and the last part gives the principal structure theorems.

The main result of this paper, Theorem 16, was announced by Chevalley in 1953, together with a proof whose basic idea was to consider the natural homomorphism from a connected algebraic group to its Albanese variety and then apply the basic properties of Albanese and Picard varieties. Chevalley's theorem also appears in the recent publications of I. Barsotti (*Ann. Mat. Pura Appl.*, vol. 38 (1955) and *Atti Accad. Naz. Lincei Rend. Cl. Sci. Fis. Mat. Nat.*, ser. 8, vol. 18 (1955)); his papers represent work done approximately simultaneously with the author's and seem to follow a similar method. We have made no use of Barsotti's papers except to appropriate from him the statement of the first half of our Proposition 2, which is not needed elsewhere.

An older version of this paper was completed before the author knew of the existence of Weil's recent papers [4], [5]; it contained roughly the

* Received December 14, 1955.

[1] During the final revision of this paper the author was connected with a project of the U. S. Air Force.

same results we present now, but in much weaker form. The main differences with the present paper are due to the fact that our original statement of Theorem 1 (which was proved by the method of Weil and Nakano) assumed that the group G contained a dense set of rational points.

Following N. Bourbaki, a map $\tau\colon V \to W$ will be called *surjective* if $\tau(V) = W$. Otherwise, the terminology and conventions we employ are for the most part those of Weil ([3], [4], [5], [6]). In particular, we follow his systematic use in [4] and [5] of the language of the Zariski topology, even for bunches of varieties. However we remark one important divergence with his usage: in this paper the closed subsets of a variety V are taken to include V itself.

The author is indebted to A. Weil, whose advice in the preparation of the final manuscript made possible many simplifications and generalizations.

1. Algebraic groups as transformation groups. We define an *algebraic group* to be the union G of a finite number of disjoint algebraic varieties $\{G_\alpha\}$ (called the *components* of G) together with a group structure on G such that for any components G_α, G_β of G, the map $g_1 \times g_2 \to g_1 g_2^{-1}$ restricted to $G_\alpha \times G_\beta$ is an everywhere defined rational map of $G_\alpha \times G_\beta$ into some component G_γ of G. A field k is called a *field of definition* of G if it is a field of definition for each component of G and for all the above rational maps $G_\alpha \times G_\beta \to G_\gamma$. If G has only one component, we say that G is *connected.*

If G is an algebraic group having k as a field of definition, consideration of the map $g \to gg^{-1} = e$ shows that the identity e of G is rational over k. It follows that the map $g \to g^{-1}$ is an everywhere defined rational map on each component of G, that the map $g_1 \times g_2 \to g_1 g_2$ is an everywhere defined rational map on each pair of components of G, and that these two maps are defined over k. If G_0 is the component of G that contains e, then $G_0 G_0^{-1} = G_0$, so G_0 is a connected algebraic group having k as a field of definition. If $g \in G$ and G_α is the component of G that contains g, then $e \in g^{-1}G_\alpha$, so $x \to g^{-1}x$ is a rational map of G_α into G_0. Similarly, $y \to gy$ is a rational map of G_0 into G_α, and it follows that G_0 and G_α are biregularly birationally equivalent, with $G_\alpha = gG_0$. Similarly $G_\alpha = G_0 g$. Thus G_0 is a normal subgroup of G and each component of G is a coset of G_0. Since G_0 is nonsingular, so is any component of G. Finally, if $g_1, g_2 \in G$, we have $k(g_1, g_2) = k(g_1, g_1 g_2) = k(g_2, g_1 g_2)$.

Let G be an algebraic group and V a variety. We say that G *operates regularly* on V (in the terminology of Weil, V is a *transformation space*

for G) if for each component G_α of G we are given an everywhere defined rational map $g \times v \to g(v)$ of $G_\alpha \times V \to V$ such that

(1) $g_1(g_2(v)) = g_1g_2(v)$ for any $g_1, g_2 \varepsilon G$, $v \varepsilon V$.

(2) $e(v) = v$ for any $v \varepsilon V$.

In this case it is clear that if k is a field of definition for G, V and the operation of G on V, and if $g \varepsilon G$, $v \varepsilon V$, then $k(g, g(v)) = k(g, v)$.

Let G be an algebraic group and V a variety. We say that G *operates* on V (or that V is a *pre-transformation space for* G) if for each component G_α of G we are given a rational map $g \times v \to g(v)$ of $G_\alpha \times V \to V$ such that if k is a field of definition for G, V, and each of these rational maps and if $g_1 \times g_2 \times v$ is a generic point over k of $G_\alpha \times G_\beta \times V$ (G_α, G_β being any components of G) then

(1) $g_1(g_2(v)) = g_1g_2(v)$.

(2) $k(g_1, g_1(v)) = k(g_1, v)$.

Consideration of the graphs of the various rational maps $G_\alpha \times V \to V$ shows this definition to be independent of the choice of the field of definition k. If G operates regularly on V, then G operates on V. When there is no danger of confusion, we shall usually write gv instead of $g(v)$. If $g_1(g_2v)$ and $(g_1g_2)v$ are both defined, then they must be equal, so we may simply write g_1g_2v.

LEMMA. *Let the algebraic group G operate on the variety V and let k be a field of definition for G, V, and the operation of G on V. If $a, b \varepsilon G$, and x is a simple point of V such that both bx and $a(bx)$ are defined, then $(ab)x$ is defined and equals $a(bx)$. If v is generic for V over $k(a)$, then av is defined and $k(a, av) = k(a, v)$; in particular av is generic for V over $k(a)$. If v is generic for V over k, then $ev = v$.*

Let g_1, g_2 be independent generic points over k of the components of G that contain a, b respectively, and let v be generic for V over $k(g_1, g_2)$. Then g_1g_2 is generic over k for the component of G that contains ab and v is generic for V over $k(g_1g_2)$. Let the point y of an affine representative of V that contains $a(bx)$ be such that $ab \times x \times y$ is a specialization over k of $g_1g_2 \times v \times (g_1g_2)v$. Then $a \times ab \times x \times y$ is a specialization over k of $g_1 \times g_1g_2 \times v \times (g_1g_2)v$. If we extend this latter specialization to a specialization over k of $g_1 \times g_2 \times g_1g_2 \times v \times g_2v \times g_1(g_2v) \times (g_1g_2)v$ we get $a \times b \times ab \times x \times bx \times a(bx) \times y$. Hence $y = a(bx)$ is uniquely determined by ab and x. It follows that $(ab)x$ is defined and equals $a(bx)$. For the rest,

we may assume that a is rational over k. If g is generic for a component of G over $k(v)$, the same is true of ga, so $(ga)v$ is defined and $k(g,(ga)v) = k(ga,(ga)v) = k(ga,v) = k(g,v)$. Thus $(ga)v$ is generic for V over $k(g) = k(g^{-1})$, so $g^{-1}((ga)v)$ is defined and generic for V over $k(g)$; thus av is defined and generic for V over $k(g)$. Since $k(g,(ga)v) = k(g,v)$, v is rational over $k(g,av)$, hence over $k(av)$; thus $k(av) = k(v)$. Finally, to show that $ev = v$, it suffices to show that $e(ev) = ev$, and this is known.

THEOREM 1. *Let the algebraic group G operate on the variety V and let k be a field of definition for G, V, and the operation of G on V. Then there exists a variety V', birationally equivalent over k to V, such that the operation of G on V' that is induced by its operation on V is regular.*

If G is connected, this is part of the main theorem of [4]. In the general case, note that G_0 operates on V, so that we may suppose that G_0 operates regularly on V. Let $\gamma_1, \cdots, \gamma_r$ be a set of points of G that are algebraic over k and such that each component of G contains at least one γ_i and let W be a k-closed proper subset of V such that $\gamma_1 v, \cdots, \gamma_r v$ are all defined whenever $v \varepsilon V - W$. If $a \varepsilon G_0$ and $v \varepsilon V - W$ then $a(\gamma_i v)$ is defined, so $(a\gamma_i)v$ is also defined, whence gv is defined for any $g \varepsilon G$ and $v \varepsilon V - W$. If $a \varepsilon G_0$, $g \varepsilon G$, $v \varepsilon V - aW$, then $a^{-1}v$ is defined and not on W, so $ga(a^{-1}v)$ is defined, whence gv is defined. Hence gv is defined for all $g \varepsilon G$, $v \not\varepsilon W' = \bigcap_a aW$, a ranging over all points of G_0. W' is clearly k-closed and $aW' = W'$ for any $a \varepsilon G_0$. Let W'' be the set of all points $p \varepsilon V$ for which there exists a $g \varepsilon G$ such that gp is defined and on W'. If $p \varepsilon W''$ we can choose $a \varepsilon G_0$ and $i = 1, \cdots, r$ such that $(a\gamma_i)p$ is defined and on W', whence $\gamma_i p$ is defined and on W'. Hence W'' is a closed proper subset of V which contains W'. Since W'' is invariant with respect to all k-automorphisms of the universal domain, W'' is k-closed. If $p \varepsilon V - W''$ and $g \varepsilon G$ then clearly $gp \not\varepsilon W''$. Hence the theorem is proved by taking $V' = V - W''$.

The following corollary give a kind of uniqueness result for Theorem 1.

COROLLARY. *Let V_1, V_2 be birationally equivalent varieties, let G be an algebraic group which operates regularly on both V_1 and V_2 in a manner consistent with their birational equivalence, and let k be a field of definition for G, V_1, V_2, the birational equivalence between V_1 and V_2, and the operation of G on V_1 and V_2. Then there exist k-closed proper subsets F_1, F_2 of V_1, V_2 respectively such that G operates regularly on both $V_1 - F_1$ and $V_2 - F_2$ and such that the birational correspondence between $V_1 - F_1$ and $V_2 - F_2$ is biregular.*

For there exist k-closed proper subsets W_1, W_2 of V_1, V_2 respectively such that the birational correspondence between $V_1 - W_1$ and $V_2 - W_2$ is biregular (for example, cf. [1, § 2, Lemma 1]) and it suffices to set

$$F_1 = \bigcap_{g \varepsilon G} gW_1, \qquad F_2 = \bigcap_{g \varepsilon G} gW_2.$$

As an example of the corollary, if our V_1, V_2 are homogeneous spaces with respect to G (i. e., G operates regularly and transitively on V_1 and V_2), then F_1 and F_2 are empty.

Let V, W be varieties and $\tau\colon V \to W$ a rational map. If k is a field of definition for V, W, τ and v is a generic point of V over k, we say that τ is *generically surjective* if τv is generic for W over k, that τ is *separable* if $k(v)$ is separably generated over $k(\tau v)$, and that τ is *purely inseparable* if $k(v)$ is a purely inseparable algebraic extension of $k(\tau v)$. Consideration of the graph of τ on $V \times W$ shows these definitions to be independent of the field of definition k.

In the following discussion leading up to Theorem 2, G will denote an algebraic group that operates on the variety V and k will denote a field of definition for G, V, and the operation of G on V. If $g \varepsilon G$ and f is a rational function on V we define another rational function $\lambda_g f$ on V by the rule $\lambda_g f(p) = f(g^{-1}p)$. Since $\lambda_{g_1}\lambda_{g_2} = \lambda_{g_1 g_2}$, the map $g \to \lambda_g$ is a homomorphism of G into the group of automorphisms of the function field on V that leave constant functions fixed. If $g \varepsilon G$ and K is an extension field of k over which g is rational, then λ_g defines an automorphism of the field $K(V)$ of all rational functions on V that are defined over K, for if v is any generic point of V over K, then $K(V)$ is naturally isomorphic to $K(v)$ and we have $K(g^{-1}v) = K(v)$.

If f is a rational function on V, we say that f is *invariant* if $\lambda_g f = f$ for all $g \varepsilon G$. Thus all constant functions on V are invariant. If F is any rational function on V, the set of all $g \varepsilon G$ such that $\lambda_g F = F$ clearly forms a closed subset of G, so to prove that F is invariant it suffices to prove that $\lambda_g F = F$ for each point g in a dense subset of G. If $F \varepsilon k(V)$ and there exist generic points $g_1, \cdots, g_n$ of the various components of G over k such that $\lambda_{g_i} F = F$ for each i, then F is invariant, for in this case $\lambda_g F = F$ for any g that is generic for a component of G over k and the set of such g's is dense in G.

We now show that if K is an extension field of k and F an invariant function in $K(V)$, then F is a rational function with coefficients in K of invariant functions in $k(V)$. To do this, it suffices to assume K a finite extension of k and hence that K is either a finite algebraic extension of k or

a simple transcendental extension of k. In the former case, let $\alpha_1, \cdots, \alpha_r$ be a basis for the vector space K over k and write $F = \alpha_1 f_1 + \cdots + \alpha_r f_r$, where each $f_i \varepsilon k(V)$. If g is generic for a component of G over K we have

$$\alpha_1(f_1 - \lambda_g f_1) + \cdots + \alpha_r(f_r - \lambda_g f_r) = 0,$$

whence, by the linear disjointness over k of K and $k(g)(V)$, $f_i - \lambda_g f_i = 0$ for each i, proving that each f_i is invariant. Finally, let $K = k(t)$, where t is transcendental over k, and write

$$F = (\sum_{i \geqq 0} f_i t^i) / (\sum_{i \geqq 0} h_i t^i),$$

where each $f_i, h_i \varepsilon k(V)$, with at least one of the f_i's or h_i's a nonzero constant, and where the polynomials $\sum f_i t^i$ and $\sum h_i t^i$ are relatively prime in the ring $k(V)[t]$. Let g be a generic point over k of a component of G such that t is transcendental over $k(g)$. Then $\sum f_i t^i$ and $\sum h_i t^i$ are relatively prime in the ring $k(g)(V)[t]$. Since

$$(\sum f_i t^i)(\sum \lambda_g h_i t^i) = (\sum h_i t^i)(\sum \lambda_g f_i t^i),$$

$\sum f_i t^i$ divides $\sum \lambda_g f_i t^i$, whence there exists $c \varepsilon k(g)(V)$ such that $\sum \lambda_g f_i t^i = c \sum f_i t^i$ and $\sum \lambda_g h_i t^i = c \sum h_i t^i$. One of the f_i's or h_i's being a nonzero constant, we have $c = 1$. Hence each f_i, h_i is invariant, proving the contention of this paragraph.

We now give a method for constructing invariant functions on V. Assume first that V is a variety in a projective space. Denote by $G_1, \cdots, G_r$ the different components of G, let v be a generic point of V over k, let g_i be a generic point of G_i over $k(v)$, $(i = 1, \cdots, r)$, and let Γ_i be the locus over $k(v)$ of the point $g_i v$. Clearly Γ_i is independent of the choice of g_i. Given any $i, j = 1, \cdots, r$, we can find a point $\alpha \varepsilon G$ that is algebraic over k and such that $g_i \alpha \varepsilon G_j$; then $g_i \alpha$ is generic for G_j over $k(v)$ and Γ_j is the locus over $k(v)$ of $(g_i \alpha) v = g_i \alpha v$. Γ_i, Γ_j are also the loci of $g_i v$, $g_i(\alpha v)$ respectively over $k(\alpha, v) = k(\alpha, \alpha v)$, and since there exists a $k(\alpha)$-automorphism of the universal domain which sends v into αv it follows that Γ_j and Γ_i have the same dimension and order. Let $F_1(v), \cdots, F_N(v) \varepsilon k(v)$ be the ratios of the Chow coordinates of the cycle $\Gamma_1 + \cdots + \Gamma_r$. If g is generic for a component of G over $k(v, g_1, \cdots, g_r)$, the loci of $g_1 g^{-1} v, \cdots, g_r g^{-1} v$ over $k(v)$ are clearly $\Gamma_1, \cdots, \Gamma_r$ in some order, whence for each $\nu = 1, \cdots, N$ we have $F_\nu(g^{-1} v) = F_\nu(v)$. Considering v to be a variable point of V, we get that each F_ν is an invariant function on V. Now all points of the bunch of varieties $\Gamma_1 \cup \cdots \cup \Gamma_r$ except possibly for those lying on a certain bunch of subvarieties of smaller dimension are of the form gv, for a suitable $g \varepsilon G$

such that gv is defined (cf. [4, Appendix, Props. 8, 10]). It follows that if v_1, v_2 are generic points of V over k such that $F_\nu(v_1)=F_\nu(v_2)$ for each ν, then there exist points g_1, g_2 that are generic for certain components of G over $k(v_1,v_2)$ such that $g_1v_1=g_2v_2$. If V is not imbedded in a projective space we get a similar result by replacing V by a projective model of V over k.

THEOREM 2. *Let the algebraic group G operate on the variety V and let k be a field of definition for G, V, and the operation of G on V. Then there exists a variety W and a generically surjective rational map $\tau\colon V\rightarrow W$, both W and τ being defined over k, which are characterized to within a birational correspondence defined over k by the following properties: $\tau\colon V\rightarrow W$ is generically surjective and separable, and if v_1, v_2 are generic points of V over k, then $\tau v_1=\tau v_2$ if and only if there exist g_1, g_2 generic for components of G over $k(v_1,v_2)$ such that $g_1v_1=g_2v_2$. If W' is any variety birationally equivalent to W, τ' the corresponding rational map from V to W', K any field of definition for V, W', and τ', and if we identify in the natural way functions on W' with functions on V, then $K(W')$ is the subfield of invariant functions of $K(V)$. If G is connected, then $K(W')$ is algebraically closed in $K(V)$.*

Let W be a variety defined over k such that $k(W)$ is k-isomorphic to the subfield of invariant functions of $k(V)$ and let τ be the natural rational map from V to W. Letting $\bar{k}$ denote the algebraic closure of k, the field $\bar{k}(W)$ is then the subfield of invariant functions of $\bar{k}(V)$. Since the points of G that are rational over $\bar{k}$ are dense in G, $\bar{k}(W)$ is the subfield of $\bar{k}(V)$ consisting of all functions left fixed by each automorphism λ_g of $\bar{k}(V)$, g ranging over the points of G that are rational over $\bar{k}$. By the first lemma of Section 3 of [2], $\bar{k}(V)$ is separably generated over $\bar{k}(W)$; hence τ is separable. In virtue of what has been done above, the first statement of the theorem is proved, except for the unicity part. So let W_1 and $\tau_1\colon V\rightarrow W_1$ have the requisite properties. Then each element of $k(W_1)$ is invariant, so $k(W_1)\subset k(W)$, and there exists a rational map $\sigma\colon W\rightarrow W_1$ defined over k such that $\tau_1=\sigma\tau$. σ is one-one for generic points of W over k, so if v is generic for V over k, then $k(\tau v)$ is purely inseparable over $k(\tau_1 v)$. Since $k(v)$ is separably generated over $k(\tau_1 v)$, we must have $k(\tau v)=k(\tau_1 v)$, proving the birationality of σ and the first part of the theorem. Now let W', τ', and K be as in the theorem. Then each function of $K(W')$ is an invariant function of $K(V)$. We thus have to show that any given invariant function of $K(V)$ is a function in $K(W')$. For this, it suffices to take K finitely generated

over the prime field, so that we may form the composed field kK. The subfield of invariant functions of $kK(V) = K(k(V))$ is $K(k(W)) = kK(W')$. Hence the invariant functions of $K(V)$ are $kK(K(W')) \cap K(V)$. But kK and $K(V)$ are linearly disjoint over K, so (e.g., using [1, § 2, Lemma 3]) the latter intersection is $K(W')$. Finally, if $F \varepsilon K(V)$ is algebraic over $K(W')$ then for any $g \varepsilon G$, $\lambda_g F$ is a conjugate of F over $K(W')$, so the closed subset of G consisting of all $g \varepsilon G$ such that $\lambda_g F = F$ is a subgroup of finite index of G. If G is connected, this subgroup must be G itself, whence $F \varepsilon K(W')$. This completes the proof.

If G operates regularly on V, then the *orbit* of a point $p \varepsilon V$ is the set of all points of the form gp, g ranging over the points of G. In this case if v_1, v_2 are generic points of V over k, the condition that $\tau v_1 = \tau v_2$ is simply that v_1 and v_2 have the same orbit. Whether or not G operates regularly on V, we call the variety W, defined to within a birational transformation, the *variety of G-orbits on V*. Strictly speaking, W is a true variety of orbits only so far as *generic* orbits are concerned.

2. Algebraic subgroups and factor groups. If G is an algebraic group and H a subgroup of G, we say that H is an *algebraic subgroup* of G if H is a closed subset of G. In this case, if k is a field of definition for G and the various components of H and if h_1, h_2 are independent generic points over k of components of H passing through e, then h_1, h_2 are each specializations over k of the point $h_1 h_2 \varepsilon H$. Thus H contains only one component H_0 passing through e and any other component of H is a coset of H_0. H is therefore an algebraic group. If H is a normal subgroup of G, so is H_0.

If G is an algebraic group, and $W_1, \cdots, W_r$ are subvarieties of G that pass through e, then the subgroup of G that is generated by the points of $W_1, \cdots, W_r$ is a connected algebraic subgroup of G; furthermore, this algebraic subgroup is defined over any field of definition for G and each W_i, and is complete if each W_i is complete. (For the easy proof, cf. [1, Theorem 6]). For example, any algebraic group G contains a *maximal* connected complete algebraic subgroup that contains all other connected complete algebraic subgroups of G. Similarly, if G is a connected algebraic group that is defined over k and if x, y are independent generic points of G over k then the locus W of $xyx^{-1}y^{-1}$ over k generates a connected algebraic subgroup G' of G, G' also being defined over k. Since any point of G' can be represented as the product of two generic point of G' over k, since any generic point of G' over k is a product of generic points of W over k, and since any generic point of W over k is a commutator of elements of G, we get that *the commu-*

tator subgroup of a connected algebraic group is a connected algebraic subgroup having the same field of definition.

By a *rational homomorphism* of an algebraic group G_1 into an algebraic group G_2 we mean a homomorphism of G_1 into G_2 which is given on each component of G_1 by an everywhere defined rational map of this component into some component of G_2. By a *biregular isomorphism* of the algebraic groups G_1 and G_2 we mean an isomorphism which is given by a set of biregular birational correspondences between the components of G_1 and those of G_2.

THEOREM 3. *Let G_1, G_2 be connected algebraic groups defined over k and let τ be a map of the generic points of G_1 over k into points of G_2 that satisfies the following conditions: if x and y are any independent generic points of G_1 over k then τx is rational over $k(x)$, $(y, \tau y)$ is a specialization over k of $(x, \tau x)$, and $\tau(xy) = \tau x \cdot \tau y$. Then τ can be extended in one and only one way to a homomorphism of G_1 into G_2, and the extended map τ is a rational map, defined over k, that is everywhere defined on G_1. If W is the locus of τx over k, then W is an algebraic subgroup of G_2 and τ is a surjective homomorphism from G_1 to W. The kernel H of τ is a k-closed normal algebraic subgroup of G_1 of dimension* $\dim G_1 - \dim W$.

COROLLARY. *Two birationally equivalent connected algebraic groups whose laws of composition correspond under the birational equivalence are biregularly isomorphic.*

This theorem and its corollary are the same as Theorem 5 and Corollary of [1], reproduced here because of their frequent future application. They can be extended without any difficulty to algebraic groups that are not connected. For example, if τ is a rational homomorphism of *any* algebraic group G_1 into an algebraic group G_2 then the kernel and image of τ are algebraic subgroups of G_1, G_2 respectively, and τG_1 is defined over any field of definition for G_1, G_2, and τ. Also, if the algebraic group H is embedded birationally and isomorphically in the algebraic group G, then H is a closed subset of G.

A suggestion of Weil is responsible for the following general proposition.

PROPOSITION 1. *Let V_1, V_2 be homogeneous spaces with respect to the connected algebraic group G and τ a rational map from V_1 to V_2 such that if g, v are independent generic points of G, V_1 respectively over some field of definition for G, V_1, V_2, the operation of G on V_1 and V_2, and τ, then $\tau(gv) = g(\tau v)$. Then τ is an everywhere defined surjective map from V_1 to*

V_2 and the relation $\tau(gv)=g(\tau v)$ holds for all $g \varepsilon G$, $v \varepsilon V_1$. If T is the graph of τ on $V_1 \times V_2$ and W_2 is any subvariety of V_2, then the cycle $W_1 = \mathrm{pr}_{V_1}((V_1 \times W_2) \cdot T)$ is defined, has dimension $(\dim V_1 - \dim V_2 + \dim W_2)$, *the point set $|W_1|$ consists precisely of all points of V_1 that are mapped into W_2 by τ, and the map induced by τ from any component of W_1 to W_2 is surjective. If τ is separable, then each component of W_1 has coefficient one and the rational map induced by τ on each component of W_1 is also separable.*

Let k be a field of definition for G, V_1, V_2, the operation of G on V_1 and V_2; and τ. Given $p \varepsilon V_1$, choose g generic for G over $k(p)$. Then $g^{-1}p$ is generic for V_1 over k, so $\tau(g^{-1}p)$ is defined, and hence also $g(\tau(g^{-1}p))$. But when v is generic for V_1 over $k(p,g)$, we have $g(\tau(g^{-1}v)) = \tau(gg^{-1}v) = \tau(v)$, so $\tau(p) = g(\tau(g^{-1}p))$ is defined. Hence $\tau(gv) = g(\tau v)$ for all $g \varepsilon G$, $v \varepsilon V_1$. By transitivity, $\tau V_1 = V_2$. T consists of precisely all points of the form $p \times \tau p$ $(p \varepsilon V_1)$, so the restriction of pr_{V_1} to T is a biregular birational correspondence between T and V_1. Since $V_1 \times V_2$ is nonsingular, to show that $(V_1 \times W_2) \cdot T$ is defined we merely have to show that the components of $(V_1 \times W_2) \cap T$ have dimension

$$\leqq (\dim V_1 \times W_2 + \dim T - \dim V_1 \times V_2) = \dim V_1 - \dim V_2 + \dim W_2.$$

But if this were not true, then for at least one point $q \varepsilon W_2$ we would have $\dim (V_1 \times q) \cap T > \dim V_1 - \dim V_2$; by transitivity, we would then have $\dim (V_1 \times q) \cap T > \dim V_1 - \dim V_2$ for *all* $q \varepsilon V_2$, which is clearly false. It follows that $(V_1 \times W_2) \cdot T$, and hence W_1, have the correct dimension, that $\tau^{-1}\{W_2\} = |W_1|$, and that $\tau|W_1| = W_2$. Now let C be any component of W_1; we wish to show that the rational map $\tau\colon C \rightarrow W_2$ is surjective. For this purpose, fix a point $p_0 \varepsilon V_1$ and consider the map $\tau_1\colon G \rightarrow V_1$ defined by $\tau_1 g = gp_0$. Letting G operate on itself by left translation, if $g_1, g_2 \varepsilon G$ we have $\tau_1(g_1g_2) = g_1g_2p_0 = g_1(\tau_1 g_2)$, so our above results apply to $\tau_1\colon G \rightarrow V_1$ and also to $\tau\tau_1\colon G \rightarrow V_2$. If C_0 is a component of the point set $\tau_1^{-1}\{C\}$, then C_0 is also a component of $(\tau\tau_1)^{-1}\{W_2\}$. The restriction of $\tau\tau_1$ to C_0 must be generically surjective to W_2, for otherwise the inverse image of a certain point of W_2 would have too large a dimension. Hence if C_0' is another component of $(\tau\tau_1)^{-1}\{W_2\}$ there exist points $g \varepsilon C_0$, $g' \varepsilon C_0'$ such that $\tau\tau_1 g = \tau\tau_1 g'$; furthermore we can suppose that C_0' is the only component of $(\tau\tau_1)^{-1}\{W_2\}$ passing through g'. Setting $\gamma = g^{-1}g'$, we get $\tau\tau_1\gamma = g^{-1}\tau\tau_1 g' = g^{-1}\tau\tau_1 g = \tau\tau_1 e$. Hence

$$\tau\tau_1 C_0\gamma = C_0\tau\tau_1\gamma = C_0\tau\tau_1 e = \tau\tau_1 C_0,$$

so $C_0\gamma \subset (\tau\tau_1)^{-1}\{W_2\}$. Since $g' \varepsilon C_0\gamma$, we have $C_0\gamma = C_0'$ and $\tau\tau_1 C_0' = \tau\tau_1 C_0$. Since all components of $(\tau\tau_1)^{-1}\{W_2\}$ have the same image under $\tau\tau_1$, we get $\tau\tau_1 C_0 = W_2$. Thus $\tau C = W_2$. It remains to prove the last two statements, so suppose τ separable. To prove that all the coefficients of W_1 are one, it suffices to prove that all the coefficients of $(V_1 \times W_2) \cdot T$ are one. Let q be a simple point of W_2. Then $V_1 \times q$ can be considered as both a $V_1 \times V_2$-cycle and a $V_1 \times W_2$-cycle, and T intersects $V_1 \times q$ properly on $V_1 \times V_2$. Furthermore, the $V_1 \times V_2$-cycle $(V_1 \times W_2) \cdot T$ is also a $V_1 \times W_2$-cycle, intersecting $V_1 \times q$ properly on $V_1 \times W_2$. By [3, Theorem 18, p. 214], any component of $(V_1 \times q) \cap T$ has the same coefficient in the cycle

$$\{(V_1 \times q) \cdot \{(V_1 \times W_2) \cdot T\}_{V_1 \times V_2}\}_{V_1 \times W_2}$$

as in the cycle $\{(V_1 \times q) \cdot T\}_{V_1 \times V_2}$. Hence to show that all the coefficients of $(V_1 \times W_2) \cdot T$ are one, it suffices to prove that each coefficient of $(V_1 \times q) \cdot T$ is one. But each $g \varepsilon G$ induces a biregular birational transformation on $V_1 \times V_2$ by the law $g(v_1 \times v_2) = gv_1 \times gv_2$ and $gT = T$. Since G is transitive on V_2, it suffices to show that each coefficient of $(V_1 \times \tau v) \cdot T$ is one, where v is generic for V_1 over k, k being as above. But $v \times \tau v$ is generic over $\overline{k(\tau v)}$ for a component of $(V_1 \times \tau v) \cdot T$, so by [3, Theorem 11, p. 161] each component of $(V_1 \times \tau v) \cdot T$ has coefficient $= [k(v) : k(\tau v)]_i = 1$. Finally, if C is a component of W_1, if the field k is also a field of definition for C and W_2, and if p is a generic point of C over k, then p is a generic point over $\overline{k(\tau p)}$ for a component of $\mathrm{pr}_{V_1}((V_1 \times \tau p) \cdot T)$. Since the latter cycle has coefficients one and is rational over $k(\tau p)$, it follows that $k(p)$ is separably generated over $k(\tau p)$. Hence the restriction of τ to C is separable.

A rational homomorphism from an algebraic group G_1 to an algebraic group G_2 is called *separable* (or *purely inseparable*) if the rational map of one component of G_1 into the corresponding component of G_2 is separable (or purely inseparable); in this case the same is true for any component of G_1.

If H is an algebraic subgroup of the algebraic group G, we shall often use the same symbol H to denote the "cycle" on G consisting of the various components of H, each taken with coefficient one. Thus if G is defined over k, we say that its algebraic subgroup H is a *rational cycle over* k if the cycle H is rational over k, i.e. if the restriction of H to each component of G is rational over k; thus H is a rational cycle over k and only if it is k-closed and all its components are defined over a separable algebraic extension of k. If H is a rational cycle over k, then (since $e \varepsilon H_0$) H_0 is left fixed by all k-automorphisms of the universal domain, so H_0 is a rational cycle over k, and thus H_0 is defined over k.

COROLLARY. *If $\tau: G \rightarrow G'$ is a surjective separable rational homomorphism from the algebraic group G to the algebraic group G', if G, G', τ are defined over k, and if H' is an algebraic subgroup of G' that is a rational cycle over k, then the algebraic subgroup H of G which is the inverse image under τ of H' is also a rational cycle over k. In particular, the kernel of τ is a rational cycle over k.*

For if G_α is any component of G, H_α' the restriction of H' to τG_α, H_α the restriction of H to G_α, and T_α is the graph of the map $\tau: G_\alpha \rightarrow \tau G_\alpha$, then we can consider G_0 to operate on both G_α and τG_α, so

$$H_\alpha = \mathrm{pr}_{G_\alpha}((G_\alpha \times H_\alpha') \cdot T_\alpha)$$

is a rational cycle over k.

We digress a bit to prove a few facts needed later. First, *if the variety V is defined over k, then the points of V that are separably algebraic over k are dense in V.* This well-known result amounts to showing that if $(x_1, \cdots, x_n)$ are quantities such that $k(x_1, \cdots, x_n)$ is separably generated over k and $F(X) \in K[X_1, \cdots, X_n]$ (K some extension field of k), $F(x) \neq 0$, then there exists a finite k-specialization $(\bar{x}_1, \cdots, \bar{x}_n)$ of $(x_1, \cdots, x_n)$ such that $F(\bar{x}) \neq 0$ and such that $k(\bar{x})$ is separably algebraic over k. Assuming $x_1, \cdots, x_r$ to be a separating transcendence basis of $k(x)$ over k and replacing k by $k(x_1, \cdots, x_{r-1})$ we reduce this to the case where $k(x)$ has transcendence degree 1 over k, in which case it is known that there exists an infinity of places of $k(x)$ over k whose residue fields are separable over k. (Our contention can also be deduced immediately from [4, Appendix, Prop. 13]). We now claim that if we have the situation described in Prop. 1, with τ separable, and if k is a field of definition for G, V_1, V_2, the operation of G on V_1 and V_2, and τ, then *for any point $p_2 \in V_2$ there exists a point $p_1 \in V_1$ such that $\tau p_1 = p_2$ and $k(p_1)$ is separably algebraic over $k(p_2)$.* For each component of the cycle $\mathrm{pr}_{V_1}((V_1 \times p_2) \cdot T)$ is defined over a separable algebraic extension of $k(p_2)$, so we may use the preceding remark.

LEMMA. *Let the algebraic group G operate on the variety V and let k be a field of definition for G, V, and the operation of G on V. Then any algebraic subgroup H of G operates on V, and if H is a rational cycle over k then the variety W of H-orbits on V and the natural rational map $\tau: V \rightarrow W$ may both be taken to be defined over k.*

By the lemma to Theorem 1, H operates on V. So let H as a cycle be rational over k, and let the overfield k' of k be a finite separable normal

algebraic extension of k over which each component of H is defined. Let $\alpha_1, \cdots, \alpha_n$ be a basis for k' over k and let $\sigma_1, \cdots, \sigma_n$ be the distinct k-automorphisms of k'. Then any H-invariant function in $k'(V)$ is of the form $\sum \alpha_i F_i$, where each $F_i \varepsilon k(V)$. Since H is a rational cycle over k, for any $j = 1, \cdots, n$, $\sum \sigma_j(\alpha_i) F_i$ is also an H-invariant function in $k'(V)$. Since $|\sigma_j(\alpha_i)| \neq 0$, each F_i is H-invariant. By Theorem 2, the lemma is proved by taking W to be any variety defined over k such that $k(W)$ is k-isomorphic to the subfield of H-invariant functions of $k(V)$, and by letting τ be the natural rational map from V to W.

THEOREM 4. *Let G be an algebraic group and H a normal algebraic subgroup of G. Then there exists an algebraic group G/H and a surjective separable rational homomorphism $\tau\colon G \to G/H$ with kernel H. These properties characterize G/H and τ to within a biregular isomorphism. Also, if k is a field of definition of G over which the cycle H is rational, then G/H and τ may be taken to be defined over k.*

This is proved by Weil [5, Prop. 2] for the case when G is connected. His proof can be extended to cover the general case, but for the sake of completeness we give a direct proof by a silghtly different method.

Let G, H, k be as above. We first prove the existence of G/H and τ in the special case where $H \subset G_0$. For any component G_α of G, G_0 operates regularly on G_α by the rule $g_0 \times g_\alpha \to g_\alpha g_0^{-1}$. Consider the induced operation of H on G_α, and construct the variety W_α of H-orbits on G_α and the natural rational map τ_α from G_α to W_α, W_α and τ_α being taken (by the lemma) to be defined over k. We wish to show first that the other obvious operation of G_0 on G_α, namely $g_0 \times g_\alpha \to g_0 g_\alpha$, induces an operation of G_0 on W_α. Let x_0 and x_α be independent generic points over k of G_0 and G_α respectively, and let $f(x_0, x_\alpha) \varepsilon k(\tau_\alpha(x_0 x_\alpha))$. Then if h is generic for a component of H over $\bar{k}(x_0, x_\alpha)$ we have $\tau_\alpha(x_0 x_\alpha h^{-1}) = \tau_\alpha(x_0 x_\alpha)$, so $f(x_0, x_\alpha h^{-1}) = f(x_0, x_\alpha)$. Imagining x_0 fixed and x_α variable on G_α, $f(x_0, x_\alpha)$ becomes an H-invariant function on G_α that is defined over $k(x_0)$, whence $f(x_0, x_\alpha) \varepsilon k(x_0, \tau_\alpha x_\alpha)$. Thus $\tau_\alpha(x_0 x_\alpha)$ is rational over $k(x_0, \tau_\alpha x_\alpha)$, so we write $x_0(\tau_\alpha x_\alpha) = \tau_\alpha(x_0 x_\alpha)$. If $x_0 \times y_0 \times x_\alpha$ is generic for $G_0 \times G_0 \times G_\alpha$ over k, then

$$x_0(y_0(\tau_\alpha x_\alpha)) = x_0(\tau_\alpha(y_0 x_\alpha)) = \tau_\alpha(x_0 y_0 x_\alpha) = x_0 y_0(\tau_\alpha x_\alpha).$$

Also, since $x_0^{-1} \times x_0 x_\alpha$ is generic for $G_0 \times G_\alpha$ over k, $\tau_\alpha x_\alpha = x_0^{-1}(\tau_\alpha(x_0 x_\alpha))$ is rational over $k(x_0, x_0(\tau_\alpha x_\alpha))$. Thus G_0 operates on W_α by the law $x_0(\tau_\alpha x_\alpha) = \tau_\alpha(x_0 x_\alpha)$, and this operation is defined over k. W_α is a prehomogeneous space with respect to G_0, since $\tau_\alpha(x_0 x_\alpha)$ is generic for W_α over $k(\tau_\alpha x_\alpha)$, so by

the main result of [4] we may, without any loss of generality, assume that W_α is a homogeneous space with respect to G_0. Then Proposition 1 shows that τ_α is defined everywhere on G_α and $\tau_\alpha G_\alpha = W_\alpha$. If $p_1, p_2 \varepsilon G_\alpha$ and x_0 is generic for G_0 over $k(p_1, p_2)$, we have $\tau_\alpha p_1 = \tau_\alpha p_2$ if and only if $\tau_\alpha(x_0 p_1) = \tau_\alpha(x_0 p_2)$, which is true if and only if $x_0 p_1$ and $x_0 p_2$ have the same H-orbit on G_α, i. e. $p_1^{-1} p_2 \varepsilon H$. Thus τ_α is a one-one map from left cosets of H on G_α to W_α. (Note that in the case $G = G_0$, we have constructed the homogeneous left coset space G/H. The normality of H has not yet been used.) Now let G_α, G_β be components of G, let x_α, x_β be independent generic points over k of G_α, G_β respectively, suppose $x_\alpha x_\beta^{-1} \varepsilon G_\gamma$, and let $f(x_\alpha, x_\beta) \varepsilon k(\tau_\gamma(x_\alpha x_\beta^{-1}))$. Imagining x_α fixed and x_β variable on G_β, f becomes an H-invariant function in $k(x_\alpha)(G_\beta)$, hence in $k(x_\alpha)(W_\beta)$; i. e. $f \varepsilon k(x_\alpha, \tau_\beta x_\beta)$. Similarly, imagining x_β fixed and x_α variable on G_α, f becomes an H-invariant function in $k(\tau_\beta x_\beta)(G_\alpha)$, and hence in $k(\tau_\beta x_\beta)(W_\alpha)$. Thus $f \varepsilon k(\tau_\alpha x_\alpha, \tau_\beta x_\beta)$. Therefore $\tau_\gamma(x_\alpha x_\beta^{-1})$ is rational over $k(\tau_\alpha x_\alpha, \tau_\beta x_\beta)$; that is, we have a rational map $\phi_{\alpha,\beta}: W_\alpha \times W_\beta \to W_\gamma$, defined by $\phi_{\alpha,\beta}(\tau_\alpha x_\alpha, \tau_\beta x_\beta) = \tau_\gamma(x_\alpha x_\beta^{-1})$. Let $g_\alpha \varepsilon G_\alpha$ and let x_0 be a generic point of G_0 over $k(g_\alpha, x_\alpha, x_\beta)$. Then the relation

$$\phi_{\alpha,\beta}(\tau_\alpha x_\alpha, \tau_\beta x_\beta) = x_0^{-1} \phi_{\alpha,\beta}(x_0(\tau_\alpha x_\alpha), \tau_\beta x_\beta)$$

implies

$$\phi_{\alpha,\beta}(\tau_\alpha g_\alpha, \tau_\beta x_\beta) = x_0^{-1} \phi_{\alpha,\beta}(x_0(\tau_\alpha g_\alpha), \tau_\beta x_\beta),$$

so $\phi_{\alpha,\beta}$ is defined whenever its second argument is generic for W_β over k. Thus if $g_\alpha \varepsilon G_\alpha$, $g_\beta \varepsilon G_\beta$, and x_0 is generic for G_0 over $k(g_\alpha, g_\beta)$, the relation

$$\phi_{\alpha,\beta}(\tau_\alpha g_\alpha, \tau_\beta g_\beta) = \phi_{\alpha,\beta}(\phi_{\alpha,0}(\tau_\alpha g_\alpha, \tau_0 x_0), \phi_{\beta,0}(\tau_\beta g_\beta, \tau_0 x_0))$$

shows that $\phi_{\alpha,\beta}$ is defined everywhere on $W_\alpha \times W_\beta$. This shows that the union $\{W_\alpha\}$ of all the W_α's (which is an abstract group in the obvious manner) is actually an algebraic group defined over k, and that the map τ, defined as τ_α on G_α, is a surjective rational homomorphism from G to $\{W_\alpha\}$ having the requisite properties. Now drop the condition $H \subset G_0$. Then $G_0 \cap H$ is a normal algebraic subgroup of G that is a rational cycle over k, so that if we first factor out $G_0 \cap H$, we are reduced to proving the existence of G/H and τ in the special case where $G_0 \cap H = \{e\}$. Here H consists of a finite number of points, each rational over k, so we may take G/H to be a set of disjoint replicas of certain components of G, exactly one component being taken from each coset of $G_0 H$, with the obvious group law on the set G/H and the obvious map τ from G to G/H. This ends the proof of the existence of G/H and τ.

For the unicity, note that if τ_1, τ_2 are surjective separable rational homo-

morphisms with kernel H from G to the algebraic groups G_1, G_2 respectively, then the unicity part of Theorem 2 (applied to the operation of $G_0 \cap H$ on G_0) shows that $(G_1)_0$ and $(G_2)_0$ are birationally equivalent. Since their group laws correspond under their birational equivalence, $(G_1)_0$ and $(G_2)_0$ are biregularly isomorphic. Hence G_1 and G_2, which are naturally isomorphic as abstract groups, are biregularly isomorphic.

Corollary 1. *G/H and τ are characterized to within a biregular isomorphism by the following properties:*

(1) *τ is a surjective rational homomorphism from G to G/H with kernel H.*

(2) *if τ' is a rational homomorphism of G into an algebraic group G' and the kernel of τ' contains H, then there exists a rational homomorphism σ of G/H into G' such that $\tau' = \sigma\tau$.*

Property (1) is known. For property (2), there exists a well-defined homomorphism $\sigma\colon G/H \to G'$ such that $\tau' = \sigma\tau$ and we need only prove σ rational. Hence we may take G connected and $G' = \tau' G$. If k is a field of definition for G, H, G/H, τ, G', τ', then $k(G')$ consists of functions in $k(G)$ that are H-invariant, so $k(G') \subset k(G/H)$. Hence σ is rational. Conversely, properties (1) and (2) clearly characterize G/H and τ to within a biregular isomorphism.

Examples. If G is an algebraic group then $G/\{e\}$ and G/G are biregularly isomorphic to G and $\{e\}$ respectively. If G and H are algebraic groups then the abstract group $G \times H$ is made into an algebraic group by identifying it with the union of the various products of components of G by components of H, and we have $(G \times H)/H$ biregularly isomorphic to G.

If H is any algebraic subgroup (not necessarily normal) of the connected algebraic group G, G/H will denote the G-homogeneous space of left cosets of H on G.

Corollary 2. *If $H \supset N$ are algebraic subgroups of the algebraic group G with N normal in G, then the natural isomorphism from H/N into G/N is biregular.*

Corollary 3. *If G is a connected algebraic group, H a connected algebraic subgroup of G, and N an algebraic subgroup of H, then the natural map from H/N into G/N is a biregular birational correspondence.*

Corollary 2 comes from Corollary 3 by restricting one's attention to the

components of the identity of G and H, so it is necessary to prove only the latter. Let $\tau\colon G \to G/N$ be the natural map. The composite of the injection of H into G and τ gives a natural everywhere defined rational map from H to τH. Letting N operate on H by the law $n(h) = hn^{-1}$, any rational function on τH induces an N-invariant function on H, so we have a natural rational map from H/N to τH. Considering G and G/N to be G-homogeneous spaces (with respect to left translation by elements of G), since τ is separable Proposition 1 shows that our map from H to τH is separable. Hence the map from H/N to τH is separable. Since the latter map is generically one-one, it is birational. Since H/N and τH are both H-homogeneous spaces (with respect to left translation by elements of H) this map from H/N to τH is biregular.

THEOREM 5. *Let G be an algebraic group that operates on the variety V and let H be a normal algebraic subgroup of G. Then G/H operates on the variety W of H-orbits on V by the rule $gH(Hv) = Hgv$ and the variety of (G/H)-orbits on W is naturally birationally equivalent to the variety of G-orbits on V.*

Let $\theta\colon G \to G/H$, $\tau\colon V \to W$ be the natural rational maps and let k be a field of definition for G, V, the operation of G on V, G/H, W, θ, and τ. Let g be generic for a component of G over k, v generic for V over $k(g)$, h generic for a component of H over $\bar{k}(g,v)$. Then $\tau(g(hv)) = \tau(ghg^{-1}(gv)) = \tau(gv)$. If we imagine g fixed and v variable we get that $\tau(gv)$ is H-invariant, so $\tau(gv)$ is rational over $k(g,\tau v)$. Also, $\tau((gh)v) = \tau(g(hv)) = \tau(gv)$, so if we imagine v fixed, g variable for a component G_α of G and h generic over $\bar{k}(v)$ for a component of $G_0 \cap H$, and if we note that θG_α is the variety of $(G_0 \cap H)$-orbits on G_α, we get that $\tau(gv)$ is rational over $k(\theta g, \tau v)$. Thus we have a generically surjective rational map from each variety $\theta G_\alpha \times W$ to W. If γ_1, γ_2 are independent generic points over k of components of G/H and w is generic for W over $k(\gamma_1,\gamma_2)$, we clearly have $\gamma_1(\gamma_2 w) = (\gamma_1\gamma_2)w$. Since $\tau v = \tau(g^{-1}(gv))$ is rational over $k(\theta g, \tau(gv))$, we get $k(\theta g, \tau(gv)) = k(\theta g, \tau v)$, so G/H operates on W. The last statement of the theorem merely says that the field of (G/H)-invariant functions on W is the field of G-invariant functions on V.

COROLLARY. *If k is a field of definition for G, V, the operation of G on V, G/H, W, the variety of (G/H)-orbits on W, the variety of G-orbits on V, and for the rational maps of G, V, W, and V respectively on the four preceding varieties, then k is a field of definition for the operation of G/H*

on W and the birational equivalence of the variety of (G/H)-orbits on W with the variety of G-orbits on V.

3. Isogeny and the homomorphism theorems. We say that the connected algebraic groups G_1 and G_2 are *isogenous* if there exists a connected algebraic group G_3 and surjective rational homomorphisms with finite kernel from G_3 to G_1 and G_2; if these rational homomorphisms are both separable, or both purely inseparable, we say that G_1 and G_2 are *separably isogenous* or *inseparably isogenous* respectively. Clearly isogenous algebraic groups have the same dimension.[2]

If one of two isogenous connected algebraic groups is commutative, then so is the other: in one direction this is trivial, in the other we use the connectedness of the commutator subgroup of a connected algebraic group.

THEOREM 6. *Isogeny is an equivalence relation among connected algebraic groups. If G and H are isogenous connected algebraic groups then there is a one-one correspondence between the connected algebraic subgroups of G and those of H such that corresponding subgroups are isogenous, and if the connected algebraic subgroups G_1 and G_2 of G correspond to the subgroups H_1 and H_2 of H, then $G_1 \supset G_2$ if and only if $H_1 \supset H_2$; if $G_1 \supset G_2$ then G_2 is normal in G_1 if and only if H_2 is normal in H_1, and in this case, G_1/G_2 and H_1/H_2 are isogenous. If G, G', H, H' are connected algebraic groups with G isogenous to H and G' isogenous to H', then $G \times G'$ is isogenous to $H \times H'$. Also, these same results hold if the notion "isogeny" is replaced throughout by either "separable isogeny" or "inseparable isogeny."*

We carry through the proof simultaneously for isogeny, separable isogeny, and inseparable isogeny. k will always denote a field of definition for all the algebraic groups and rational homomorphisms in question at any time. Obviously the relations isogeny, etc., are each reflexive and symmetric. To prove transitivity, let σ_1, σ_2 be surjective rational homomorphisms with finite

[2] In [6] two abelian varieties A and B are said to be isogenous if they have the same dimension and if there exists a surjective rational homomorphism from A to B. This is an equivalence relation for abelian varieties but not for more general groups, which accounts for the necessity of our present definition (clearly an extension of the older one). For an example in which the old definition is inadequate, let G_1 be the algebraic group of all $n \times n$ matrices of determinant 1, where $n > 1$ is not a power of the field characteristic, and let $G_2 = G_1/C$, C being the center of G_1. C consists of all multiples of the unit matrix by n-th roots of unity, hence has finite order > 1, while G_2 is the projective group. Since G_2 has no proper normal subgroups it possesses no surjective homomorphism to G_1.

kernel from the connected algebraic group G to the algebraic groups G_1, G_2 respectively, and let τ_2, τ_3 be surjective rational homomorphisms with finite kernel from the connected algebraic group G' to the algebraic groups G_2, G_3 respectively. Let $\sigma_2 \times \tau_2$ be the obvious surjective homomorphism from $G \times G'$ to $G_2 \times G_2$ and let Γ be the component of the identity of the inverse image under $\sigma_2 \times \tau_2$ of the diagonal on $G_2 \times G_2$. Since $\sigma_2 \times \tau_2$ has finite kernel, Γ has the same dimension as G_2. If $p \times p'$ is a generic point of Γ over k, then p, p' are generic over k for G, G' respectively, $\sigma_2 p = \tau_2 p'$, and $k(p \times p')$ is algebraic over $k(\sigma_2 p)$. The maps $p \times p' \to p$, $p \times p' \to p'$ define surjective rational homomorphisms with finite kernel from Γ to G, G' respectively, and hence we have surjective rational homomorphisms with finite kernel from Γ to both G_1 and G_3. In the case of separable (or inseparable) isogeny, $k(p \times p')$ is separable (or purely inseparable) over $k(\sigma_2 p)$ and hence $k(p \times p')$ is separable (or purely inseparable) over each of the fields $k(p)$ and $k(p')$. Thus the relations isogeny, separable isogeny, and inseparable isogeny are equivalence relations. Next let σ, τ be surjective rational homomorphisms with finite kernel from the connected algebraic group Γ to the algebraic groups G, H respectively. We say that the connected algebraic subgroups G_1, H_1 of G, H respectively correspond if there exists a connected algebraic subgroup Γ_1 of Γ such that $\sigma\Gamma_1 = G_1$ and $\tau\Gamma_1 = H_1$. Since for any connected algebraic subgroup G_1 of G there exists one and only one connected algebraic subgroup Γ_1 of Γ such that $\sigma\Gamma_1 = G_1$ (namely $\Gamma_1 =$ component of the identity of $\sigma^{-1}(G_1)$), we get the one-one correspondence claimed by the theorem, and the corresponding groups G_1, H_1 are clearly isogenous. In the case of separable isogeny we can apply Proposition 1 to the separable map $\sigma\colon \Gamma \to G$; this shows that the rational homomorphism from Γ_1 to G_1 is separable, so corresponding subgroups G_1, H_1 are separable isogenous. In the case of inseparable isogeny, σ and τ are isomorphisms, so the homomorphisms from Γ_1 to G_1 and H_1 are isomorphisms, hence purely inseparable. If G_1, G_2 correspond to H_1, H_2, then clearly $G_1 \supset G_2$ if and only if $H_1 \supset H_2$. For the statements about normality and factor groups, it suffices to take $G_1 = G$, $H_1 = H$. Then if G_2 is normal in G, $\sigma^{-1}(G_2)$ is normal in Γ, so the component of the identity of $\sigma^{-1}(G_2)$ is normal in Γ, whence H_2 is normal in H. If Γ_1 is a connected normal algebraic subgroup of Γ and $G_1 = \sigma\Gamma_1$, $H_1 = \tau\Gamma_1$, then we have surjective rational homomorphisms from Γ to G/G_1 and H/H_1. The kernels of these homomorphisms contain Γ_1, so we get surjective rational homomorphisms from Γ/Γ_1 to G/G_1 and H/H_1. The kernels of these last two homomorphisms being finite, G/G_1 and H/H_1 are isogenous. In the case of separable isogeny, all the above homomorphisms

are separable, so G/G_1 and H/H_1 are separably isogenous. In the case of inseparable isogeny, the homomorphisms from Γ/Γ_1 to G/G_1 and H/H_1 are isomorphisms, so G/G_1 and H/H_1 are inseparably isogenous. The statement about direct products is immediate.

COROLLARY 1. *If G and H are isogenous connected algebraic groups and if $G = G_0 \supset G_1 \supset G_2 \supset \cdots$ is a normal chain of connected algebraic subgroups of G, then there exists a normal chain of connected algebraic subgroups of H, say $H = H_0 \supset H_1 \supset H_2 \supset \cdots$ such that each G_i is isogenous to H_i and each G_i/G_{i+1} is isogenous to H_i/H_{i+1}. The same result holds if we replace "isogeny" by either "separable isogeny" or "inseparable isogeny."*

If G_1 and G_2 are algebraic groups (not necessarily connected), we say that G_1 and G_2 are *inseparably isogenous* if there exists an algebraic group G_3 and surjective rational isomorphisms from G_3 to G_1 and G_2. This agrees with the previous definition in the case of connected algebraic groups. The following result is got by trivially modifying the above proofs.

COROLLARY 2. *Theorem 6 and Corollary 1 hold for inseparable isogeny if we delete the word "connected."*

In characteristic zero inseparable isogeny is the same as biregular isomorphism. In characteristic $p \neq 0$, consider the automorphism Θ of the universal domain defined by $\Theta\alpha = \alpha^p$. Under this automorphism a point in affine space with coordinates $(\alpha_1, \cdots, \alpha_r)$ will go into the point $(\alpha_1^p, \cdots, \alpha_r^p)$, a variety V will go into a variety ΘV, and an algebraic group G will go into an inseparably isogenous algebraic group ΘG. If τ is a surjective rational isomorphism from the algebraic group G to the algebraic group G', then for n sufficiently large the isomorphism $\Theta^n\tau^{-1}$ from G' to $\Theta^n G$ will be rational. Hence if G_1, G_2 are inseparably isogenous algebraic groups the obvious surjective isomorphism from G_1 to $\Theta^n G_2$ will be rational for n sufficiently large.

THEOREM 7. *If $H \supset N$ are algebraic subgroups of the connected algebraic group G with N normal in G, then the natural one-one correspondence between the points of G/H and $(G/N)/(H/N)$ is birational and biregular. If G is not necessarily connected but both H and N are normal in G, then the natural isomorphism between G/H and $(G/N)/(H/N)$ is biregular.*

The second part of the theorem follows from the first by restricting one's attention to G_0, $G_0 \cap H$ and $G_0 \cap N$. To prove the first part, note first that we have a sequence of natural separable rational maps

$$G \to G/N \to (G/N)/(H/N),$$

so the natural map from G to $(G/N)/(H/N)$ is a separable rational map. Bearing in mind that G/H is the variety of H-orbits on G (H operating on G by the law $h(g)=gh^{-1}$), and similarly for G/N, H/N, and $(G/N)/(H/N)$, we see that any rational function on $(G/N)/(H/N)$ gives rise to an H/N-invariant function on G/N, and thence to an H-invariant function on G. Thus there exists a natural generically surjective rational map from G/H to $(G/N)/(H/N)$. This last map being separable and generically one-one, it is actually birational. Since both G/H and $(G/N)/(H/N)$ are G-homogeneous spaces (G operating by left translation), the birational correspondence between G/H and $(G/N)/(H/N)$ is biregular.

THEOREM 8. *Let H and N be algebraic subgroups of the algebraic group G, with N normal in G. Then HN is an algebraic subgroup of G and the natural surjective isomorphism from $H/(H\cap N)$ to HN/N is rational.*

By the comments at the beginning of Section 2, the subgroup H_0N_0 of G is algebraic; hence HN is an algebraic subgroup of G. The natural isomorphism from H into HN is rational, so the homomorphism from H to HN/N is rational. The kernel of this homomorphism being $H\cap N$, we have a rational isomorphism from $H/(H\cap N)$ to HN/N.

If V is a variety and $\tau\colon V\to W$ a generically surjective rational map, k a field of definition for V, W, and τ, and x a generic point of V over k, the *order of inseparability of* τ is $[k(x):k(\tau x)]_i$; this is independent of the choice of x and, by [3, Theorem 11, p. 161], independent of the field k. If τ is a surjective rational homomorphism from an algebraic group G to an algebraic group G', by the *order of inseparability of* τ we mean the order of inseparability of the restriction of τ to any component of G; this is independent of the component chosen.

The following proposition, which gives more precise information on Theorem 8, will not be used in the sequel.

PROPOSITION 2. *Let H, N be algebraic subgroups of the connected algebraic group G, with N normal in G and $HN=G$. Then H and N intersect properly on G and $H\cdot N=q(H\cap N)$, where q is the order of inseparability of the natural rational isomorphism from $H/(H\cap N)$ to HN/N. If k is a field of definition for H, N and G, and h and n are independent generic points of components of H and N respectively over k, then $q=[k(h,n):k(hn)]_i$.*

Let τ_1 be the homomorphism from H to $H/(H\cap N)$ and τ the homomorphism from G to $HN/N=G/N$. Then if k, h, n are as above, the

order of inseparability q of the homomorphism from $H/(H \cap N)$ to HN/N is given by $q = [k(\tau_1 h) : k(\tau h)]_\iota$. Since τ_1 is separable, [3, Prop. 27, p. 23] gives $q = [k(h) : k(\tau h)]_\iota$. Using the paragraph after Prop. 1, Cor., choose $\bar{h} \varepsilon G$ such that $\tau\bar{h} = \tau h$ and $k(\bar{h})$ is a separable algebraic extension of $k(\tau h)$. By [3, Prop. 25, p. 22], $q = [k(h, \bar{h}) : k(\bar{h})]_\iota$. The point $\bar{h}^{-1}hn$ is generic for a component of N over $k(h, \bar{h})$, $k(\bar{h}, \bar{h}^{-1}hn)$ is a regular extension of $k(\bar{h})$, and the fields $k(h, \bar{h})$ and $k(\bar{h}, \bar{h}^{-1}hn)$ are free with respect to each other over $k(\bar{h})$, hence linearly disjoint over $k(\bar{h})$. By [3, Prop. 26, p. 23], $q = [k(h, \bar{h}, \bar{h}^{-1}hn) : k(\bar{h}, \bar{h}^{-1}hn)]_\iota = [k(h, n, \bar{h}) : k(hn, \bar{h})]_\iota$. Since $k(\bar{h})$ is separably algebraic over $k(\tau h) = k(\tau(hn))$, [3, Proposition 25, p. 22] gives $q = [k(h, n) : k(hn)]_\iota$. Next, the equality of the dimensions of $H/(H \cap N)$ and G/N implies that the cycles H and N intersect properly on G. Since there exists a translation on G taking any given component of $H \cap N$ into any other given component, these components all occur to the same multiplicity in $H \cdot N$. We shall complete the proof by showing that this common multipilicity is $[k(h, n) : k(hn)]_\iota$. But we could have taken h, n to lie in H_0, N_0 respectively, so from now on we may assume that H and N are *connected.* If $p \varepsilon G$ we can write $p = h_1 n_1$, with $h_1 \varepsilon H$, $n_1 \varepsilon N$, and therefore $H \cap pN = H \cap h_1 N = h_1(H \cap N)$. Thus H and pN intersect properly on G and the common multiplicity of the components of $H \cap N$ in $H \cdot N$ equals the multiplicity of any component of $H \cap pN$ in $H \cdot pN$. But by [6, Cor. 1, p. 24], $p \times H \cdot pN = W \cdot (p \times G)$, where $W \subset G \times G$ is the locus over k of $hn^{-1} \times h$. W is also the locus over k of $hn \times h$. Taking p to be the point hn and applying [3, Theorem 11, p. 161], we get that each component of $W \cdot (p \times G)$ has multiplicity $[k(hn, h) : k(hn)]_\iota = [k(h, n) : k(hn)]_\iota$. This completes the proof.

COROLLARY. *Let H, N be connected algebraic subgroups of the algebraic group G and suppose that almost all points of G are of the form hn, with $h \varepsilon H$ and $n \varepsilon N$. Then the map $\tau\colon H \times N \to G$ defined by $\tau(h \times n) = hn$ is birational if and only if $H \cdot N = e$. If N is a normal subgroup of G and $H \cdot N = e$, then τ is biregular.*

Let k be a field of definition for H, N, and G and let $h \times n$ be a generic point of $H \times N$ over k. Then hn is generic for G over k. Note that the last part of the proof of the proposition made no use of the normality of N, and in fact showed that if H and N intersect properly on G, then $H \cdot N = [k(h, n) : k(hn)]_\iota (H \cap N)$. If τ is birational, it must be one-one almost everywhere, so $H \cap N = \{e\}$. Furthermore, in this case $\dim G = \dim H + \dim N$, so $H \cdot N$ is defined. Since $k(hn) = k(h, n)$, we have

$H \cdot N = e$. Conversely, let $H \cdot N = e$. Since τ is one-one almost everywhere, $k(h, n)$ is purely inseparable over $k(hn)$. Since $[k(h,n) : k(hn)]_\iota = 1$, we have $k(h,n) = k(hn)$, so τ is birational. If in addition N is normal in G then τ is one-one and surjective, so Zariski's main theorem on birational correspondences gives the biregularity of τ.

We now prove the analogues for algebraic groups of the Zassenhaus lemma and the Jordan-Hölder-Schreier theorem.

LEMMA. *Let $H_1 \supset N_1$ and $H_2 \supset N_2$ be algebraic subgroups of the algebraic group G, with N_1, N_2 normal in H_1, H_2 respectively. Then $N_1(H_1 \cap N_2)$ and $N_2(H_2 \cap N_1)$ are normal algebraic subgroups of the algebraic groups $N_1(H_1 \cap H_2)$ and $N_2(H_2 \cap H_1)$ respectively and the algebraic groups*

$$\frac{N_1(H_1 \cap H_2)}{N_1(H_1 \cap N_2)} \quad \textit{and} \quad \frac{N_2(H_2 \cap H_1)}{N_2(H_2 \cap N_1)}$$

are naturally isomorphic, the isomorphism being an inseparable isogeny.

By Theorem 8, $N_1(H_1 \cap N_2)$ and $N_1(H_1 \cap H_2)$ are algebraic subgroups of H_1, and the first is normal in the second. There exists a natural rational isomorphism from $H_1 \cap H_2$ into $N_1(H_1 \cap H_2)$ and hence there exists a natural surjective rational homomorphism from $H_1 \cap H_2$ to

$$N_1(H_1 \cap H_2)/N_1(H_1 \cap N_2).$$

But the kernel of this last homomorphism is

$$(H_1 \cap H_2) \cap (N_1(H_1 \cap N_2)) = (H_1 \cap N_2)(H_2 \cap N_1).$$

Hence there exists a natural surjective rational isomorphism from

$$(H_1 \cap H_2)/(H_1 \cap N_2)(H_2 \cap N_1) \text{ to } N_1(H_1 \cap H_2)/N_1(H_1 \cap N_2).$$

The whole lemma now follows from symmetry.

THEOREM 9. *If G is an algebraic group, then any two normal chains of algebraic subgroups of G have refinements consisting of normal chains of algebraic subgroups whose successive algebraic factor groups, except for order, are inseparably isogenous.*

For if

$$G = G_0 \supset G_1 \supset \cdots \supset G_r = \{e\} \text{ and } G = H_0 \supset H_1 \supset \cdots \supset H_s = \{e\}$$

are the given normal chains, the standard proof of the algebraic version of this theorem (cf. Zassenhaus, *Lehrbuch der Gruppentheorie*) consists in inter-

posing in the first chain groups of the form $G_i(G_{i-1} \cap H_j)$ and analogous groups in the second chain, so it all reduces to the lemma.

If $G = G_0 \supset G_1 \supset \cdots \supset G_r = \{e\}$ is a normal chain of algebraic subgroups of the algebraic group G, and if H_i is the component of the identity of G_i $(i = 0, \cdots, r)$, then $H_0 \supset H_1 \supset \cdots \supset H_r = \{e\}$ is a normal chain of algebraic subgroups of H_0 and each H_{i-1}/H_i $(i = 1, \cdots, r)$ is separably isogenous to the component of the identity of G_{i-1}/G_i. Define an algebraic group to be *simple* if it contains no proper connected normal algebraic subgroup. The following results immediately from Theorem 9 and Theorem 6, Cor. 1.

COROLLARY. *If G and H are isogenous connected algebraic groups and if $G = G_0 \supset G_1 \supset \cdots \supset G_r = \{e\}$ and $H = H_0 \supset H_1 \supset \cdots \supset H_s = \{e\}$ are normal chains of connected algebraic subgroups such that each G_{i-1}/G_i and each H_{i-1}/H_i is simple and of dimension > 0, then $r = s$ and the successive algebraic factor groups of G and H are isogenous, except for order.*

4. Solvable algebraic groups and cross sections. Throughout the remainder of this paper G_a will denote the algebraic group consisting of the affine line with the composition law $(x_1)(x_2) = (x_1 + x_2)$ and G_m will denote the algebraic group of the affine line minus the origin with the composition law $(x_1)(x_2) = (x_1 x_2)$. Any connected noncomplete algebraic group of dimension one is biregularly isomorphic to either G_a or G_m. If p is the field characteristic, then G_m has elements of any given finite order not divisible by p and no element of order p, while each element $\neq e$ of G_a has order p if $p \neq 0$ and e is the only element of G_a having finite order if $p = 0$. A rational image of a rational curve being rational, any rational homomorphic image with finite kernel of G_a or G_m is biregularly isomorphic to G_a or G_m respectively. Using the fact that if an algebraic group G has a noncomplete rational homomorphic image then G itself is noncomplete, we get that any connected algebraic group that is isogenous to G_a or G_m is biregularly isomorphic to G_a or G_m respectively, while G_a and G_m are themselves nonisogenous.

We say that an algebraic group is *solvable* if it has a normal chain of algebraic subgroups such that each algebraic factor group is biregularly isomorphic to either G_a, G_m, or a finite commutative group. If one of two inseparably isogenous algebraic groups is solvable, so is the other. A connected solvable algebraic group has a normal chain of connected algebraic subgroups such that each algebraic factor group is biregularly isomorphic to G_a or G_m, hence if one of two isogenous connected algebraic groups is solvable,

so is the other. By standard arguments, any algebraic subgroup of a solvable algebraic group is solvable, any rational homomorphic image of a solvable algebraic group is solvable, and if an algebraic group G contains a normal solvable algebraic subgroup H such that G/H is solvable, then G is solvable.

An *algebraic group of matrices* is an algebraic subgroup of the multiplicative group of all invertible square matrices (of some degree $n > 0$) with coefficients in the universal domain. For example, the full triangular group T of degree n consisting of all invertible $n \times n$ matrices (a_{ij}) such that $a_{ij} = 0$ if $i > j$, and its subgroups T_ν $(\nu = 1, \cdots, n)$ defined by $a_{ij} = 0$ if $i > j$ or if $1 \leqq j - i < \nu$, while $a_{11} = \cdots = a_{nn} = 1$, are all connected algebraic groups of matrices, and $T \supset T_1 \supset \cdots \supset T_n = \{e\}$, where e is the unit matrix. The map $(a_{ij}) \rightarrow (a_{11}, \cdots, a_{nn})$ clearly is a surjective separable rational homomorphism from T to $(G_m)^n$ with kernel T_1, so T_1 is normal in T and $T/T_1 = (G_m)^n$. Similarly, if $1 \leqq \nu < n$, the map $(a_{ij}) \rightarrow (a_{1,\nu+1}, a_{2,\nu+2}, \cdots, a_{n-\nu,n})$ is a surjective separable rational homomorphism from T_ν to $(G_a)^{n-\nu}$ with kernel $T_{\nu+1}$, so $T_{\nu+1}$ is normal in T_ν and $T_\nu/T_{\nu+1} = (G_a)^{n-\nu}$. Thus the full triangular group T is solvable.

Now let $S = \{s\}$ be any set of $n \times n$ matrices which commute with each other. Considering S as a set of linear transformations on an n-dimensional vector space, for any $s \varepsilon S$ and any quantity α, the null space of $(s - \alpha e)$ is S-invariant. Hence the set of matrices S is reducible (in the sense of representation theory), and by repeated application of this we can reduce all the matrices of S simultaneously to triangular form. Hence every commutative algebraic group of matrices is solvable.

Let the algebraic group G operate on the variety V and let W be the variety of G-orbits on V and τ the natural rational map from V to W. A rational map σ from W into V is called a *cross section* if $\tau\sigma = 1$; this of course implies that if k is a field of definition for G, V, the operation of G on V, W, τ, and σ, and if p is a generic point of W over k, then σp is a point of V at which τ is defined. (Note however that σ need not be defined at all points of W, so σ is a cross section in the topological sense only on an open subset of W.) The most important case is that in which G is connected and V is a principal space with respect to G, i.e., k being as above and $g \times v$ being generic for $G \times V$ over k, then G operates regularly on V and there exists an everywhere defined rational map (defined over k) from the locus over k of $v \times gv$ into G such that $v \times gv \rightarrow g$. (For example, V may be a connected algebraic group and G a connected algebraic subgroup of V operating on V by either of the laws $g(v) = gv$ or $g(v) = vg^{-1}$; W may be taken to be the corresponding homogeneous space.) In this case $p = \tau v$ is

generic for W over k, so σp is defined and hence $k(g) \subset k(\sigma p, g(\sigma p))$. Since $\tau(g(\sigma p)) = p$, we get $k(g(\sigma p)) = k(g, p)$. Clearly the transcendence degree of $k(g, p)$ over k is $\dim G + \dim W = \dim V$, so $g(\sigma p)$ is generic for V over k. Thus V is birationally equivalent over k to $G \times W$, and G operates only on the first factor, there by multiplication on the left.

In the proof of the following lemma we use a number of results in the theory of algebraic curves which may be found in Chevalley's *Introduction to the Theory of Algebraic Functions of One Variable.*

LEMMA. *Let V be a homogeneous space with respect to the algebraic group G $(= G_a$ or $G_m)$ and let k be a field of definition for V and the operation of G on V. Then V has a point that is rational over k.*

If g, p are independent generic points over k of G, V respectively, then gp is generic for V over $k(p)$. In particular, $\dim V \leqq 1$. The case $\dim V = 0$ is trivial, so suppose V is a curve. Since $k(p, gp)$ is a subfield of the purely transcendental extension $k(p, g)$ of $k(p)$, it is an algebraic function field of one variable over the constant field $k(p)$ that has genus zero. But $k(p, gp)$ may be got by taking the function field $k(gp)/k$ and extending the constant field from k to $k(p)$, which is a regular extension of k. Hence $k(gp)/k$ has genus zero; i. e. $k(V)/k$ has genus zero. Therefore V is birationally equivalent over k to a conic C lying in the projective plane. C is complete and both C and V are (absolutely) nonsingular, so the birational correspondence between V and C is biregular between V and $C - S$, where S is a finite subset of C. Thus we may suppose that $V = C - S$. Now any birational transformation on C, in particular one induced by an element of G, has at least one fixed point. Since any nontrivial translation of a homogeneous space with respect to a commutative group is free of fixed points, this implies that S is nonempty. Let D be the projective line containing G $(D = G_a \cup (\infty)$ or $G_m \cup (0) \cup (\infty))$. Then there is an everywhere defined surjective rational map ϕ_p from D to C such that ϕ_p is defined over $k(p)$ and $\phi_p g = gp$. Clearly $S \subset \phi_p(D - G)$. Since the points of S are rational over $k(p)$ and since p could have been taken to be any generic point of V over k, the points of S are rational over k. Since C has genus zero and at least one rational point, it is birationally equivalent over k to a projective line. C thus contains at least three rational points, and since S has at most two, V has a rational point.

Note that the lemma is false if we merely assume that G, an algebraic group defined over k, is biregularly isomorphic (over some extension field of k) to G_a or G_m. For example, let k be the real numbers and let G be the

rotation group in two dimensions operating in the obvious manner on the finite part of the curve $x^2+y^2=-1$.

If G is a connected algebraic group possessing a normal chain of connected algebraic subgroups $G=G_0\supset G_1\supset\cdots\supset G_r=\{e\}$ such that each G_i $(i=0,\cdots,r-1)$ is defined over k and possesses a surjective separable rational homomorphism with kernel G_{i+1} to G_a or G_m, the homomorphism also being defined over k, we say that k is a *field of definition for the solvability of* G.

THEOREM 10. *If the connected solvable algebraic group G operates regularly on the variety V and if $\tau\colon V\to W$ is the natural rational map from V to the variety of G-orbits on V, then there exists a cross section $\sigma\colon W\to V$. Furthermore, if k is a field of definition for the solvability of G, for V, the operation of G on V, W, and τ, then σ may be taken to be defined over k.*

In the special case where V is a homogeneous space with respect to G, the theorem merely says that V has a point that is rational over k. We first assume this special case and show that the theorem holds generally. So let G, V, W, τ, k be as above and let v be a generic point of V over k. Then $p=\tau v$ is generic for W over k. By Theorem 2, $k(v)$ is separably generated over $k(p)$ and, since G is connected, $k(p)$ is algebraically closed in $k(v)$. Thus $k(v)$ is a regular extension of $k(p)$, and the locus V_1 of v over $k(p)$ is a variety defined over $k(p)$ that is a closed subset of V. If g is generic for G over $k(v)$, then gv is generic for V over k and $\tau(gv)=p$. Thus the natural k-isomorphism between $k(v)$ and $k(gv)$ is actually a $k(p)$-isomorphism, and so $gv\in V_1$. It follows that G operates regularly on V_1. If v' is any generic point of V_1 over $k(p)$ we have $k(p)(v')$ naturally $k(p)$-isomorphic to $k(p)(v)$, so $k(v')$ is naturally k-isomorphic to $k(v)$; thus v' is a generic point of V over k. Since $\tau v'=\tau v$, Theorem 2 shows that there exists $g'\in G$ such that $g'v=v'$. Since v' could have been taken to be generic for V_1 over $k(v)$, this implies that gv is generic for V_1 over $k(v)$, i.e. V_1 is a prehomogeneous space with respect to G. Thus there exists a homogeneous space birationally equivalent to V_1 over $k(p)$. Since G operates regularly on V_1, the corollary to Theorem 1 shows that this homogeneous space may be taken to be a $k(p)$-open subset V_1' of V_1. Thus there exists a point $p_1\in V_1'$ that is rational over $k(p)$. Choose $\gamma\in G$ such that $p_1=\gamma v$. Since $\tau=\tau\gamma^{-1}$, τ is defined at p_1 and $\tau p_1=\tau v=p$. Thus $\sigma\colon p\to p_1$ gives a cross section that is defined over k. We must now prove our original assumption that if V is homogeneous, then V has a point that is rational over k. Since this is trivial

if $\dim G = 0$ and the lemma if $\dim G = 1$, we assume that $\dim G > 1$ and use induction on $\dim G$. Let G_1 be a connected algebraic subgroup of G having k as a field of definition for its solvability and such that G admits a surjective separable rational homomorphism defined over k and with kernel G_1 to G_a or G_m. Let V' be the variety of G_1-orbits on V and τ' the natural rational map from V to V', both V' and τ' being taken to be defined over k. By Theorem 5 and Corollary, G/G_1 operates on V', this operation being defined over k. If g, v are independent generic points of G, V respectively over k and if ϕ is the natural homomorphism from G to G/G_1, we have $\tau'(gv) = \phi g(\tau' v)$. Since gv is generic for V over $k(v)$, $\phi g(\tau' v)$ is generic for V' over $k(\tau' v)$; i. e. V' is a prehomogeneous space with respect to G/G_1. Thus we may suppose that V' is homogeneous with respect to G/G_1. Then there exists a point $p' \varepsilon V'$ that is rational over k. But V' may also be considered as a homogeneous space with respect to G, so Proposition 1 is applicable to the case $\tau'\colon V \to V'$. Let S be the set of all points of V whose τ'-image is p'. Then S is a closed subset of V which, considered as a cycle all of whose coefficients are one, is rational over k. But if $v_1, v_2 \varepsilon V$, then $\tau' v_1 = \tau' v_2$ if and only if $v_2 \varepsilon G_1 v_1$. (This last relation holds when v_1, v_2 are generic for V over k, hence, by transitivity, for all v_1, v_2.) Thus S has only one component, and hence is a variety defined over k. Clearly S is a homogeneous space with respect to G_1. By the induction hypothesis, S, and hence V, has a point that is rational over k.

COROLLARY 1. *If H is a connected solvable algebraic subgroup of the connected algebraic group G, then G is birationally equivalent to $H \times (G/H)$. Furthermore, this birational equivalence may be defined over any field of definition for G, the solvability of H, and the map $G \to G/H$.*

This follows from the discussion immediately preceding the lemma. Note that if $\sigma\colon G/H \to G$ is a cross section for the action of H on G, then the birational equivalence is given by $h \times gH \leftrightarrow \sigma(gH)h^{-1}$ or $h \times Hg \leftrightarrow h\sigma(Hg)$ according as G/H is the homogeneous space of left or right cosets of G modulo H.

COROLLARY 2. *If G is a connected solvable algebraic group, then G is rational. More precisely, if k is a field of definition for the solvability of G, then $k(G)$ is a purely transcendental extension of k.*

5. The main structure theorems. An *abelian variety* is a complete connected algebraic group. The basic facts about abelian varieties can be

found in [6]. Since abelian varieties are commutative groups, when there is no danger of confusion we shall denote their group operations additively. The word "abelian" will always refer to abelian varieties, rather than to commutativity.

By a *linear group* we mean an algebraic group which is biregularly isomorphic to an algebraic group of matrices. Note that linear groups, being embeddible in affine spaces, are noncomplete if they have dimension > 0; thus $\{e\}$ is the only algebraic group which is both linear and abelian. Algebraic subgroups and direct products of linear groups (or abelian varieties) are again linear groups (or complete algebraic groups). The groups G_a and G_m are both linear, the former possessing the matrix representation

$$x \to \begin{pmatrix} 1 & x \\ 0 & 1 \end{pmatrix}.$$

If G is an algebraic group and $g \varepsilon G$, we denote the normalizer of g (i.e. the set of all elements of G that commute with g) by N_g. This is clearly an algebraic subgroup of G. The center C of G is the intersection of all N_g's so C is also an algebraic subgroup of G. The component of the identity of an algebraic group G will always be denoted G_0.

THEOREM 11. *Any rational homomorphism of an abelian variety into a linear group, or of a connected linear group into an abelian variety, is trivial.*

Since a rational homomorphic image of an abelian variety is abelian, the first part is proved. So let τ be a rational homomorphism of the connected linear group L into the abelian variety A. τ is trivial if $\dim L = 0$, so suppose $\dim L > 0$ and that our proposition is true for all linear groups of smaller dimension. If L is commutative then according to Section 4 it is solvable, hence rational, hence $\tau L = 0$. If L is not commutative, let k be a field of definition for L, A, and τ and let g be generic for L over k. Then $\dim N_g < \dim L$, so $\tau(N_g)_0 = 0$. Since $g \varepsilon N_g$, τg has some finite order ν. But τg is generic for τL over k, so for each point $p \varepsilon \tau L$ we have $\nu \cdot p = 0$. Since τL is abelian, $\tau L = 0$.

If G is an algebraic group that operates on the variety V, f a rational function on V, and $g \varepsilon G$, we define (as before) the function $\lambda_g f$ by $\lambda_g f(p) = f(g^{-1}p)$. In the following, we consider vector spaces of functions on V over constants, i.e. over the universal domain. In particular, we consider finite dimensional vector spaces S of functions on V with the property that $\lambda_g S = S$ for all $g \varepsilon G$. If V is a homogeneous space with respect to G

then any such space S consists of functions that are finite everywhere on V, for if W is a proper closed subset of V such that each function of S is finite on $V-W$, then each function in $S=\lambda_g S$ is finite on $V-gW$ for all $g \varepsilon G$. We need another definition: If the algebraic group G operates on V, there may exist a surjective rational homomorphism τ from G to an algebraic group G' that also operates on V, and in such a way that $gp=(\tau g)p$; if any such τ is necessarily a biregular isomorphism, we say that G operates *faithfully* on V. By [4, Proposition 2], this is equivalent to the definition given by Weil when G is connected. Clearly the operation of any connected algebraic group on itself by left translation is faithful.

THEOREM 12. *Let the algebraic group G operate regularly on the non-singular variety V, all defined over the field k. Then any everywhere finite rational function on V is contained in a finite dimensional vector space S of such functions such that $\lambda_g S=S$ for all $g \varepsilon G$, and S may be taken so as to have a basis consisting of functions that are defined over k. If λ_g induces the linear transformation Λ_g on such a space S, then (choosing such a basis for S) the map $g \rightarrow \Lambda_g$ is a surjective rational homomorphism defined over k from G to an algebraic group of matrices. If G operates faithfully on V and the functions of S generate the entire function field of V, then this homomorphism is a biregular isomorphism.*

By [3, Theorem 10, p. 239], any everywhere finite rational function f on V can be written $f=\sum_\nu \alpha_\nu \phi_\nu$, where the α_ν's are constants and each ϕ_ν is an everywhere finite function in $k(V)$, so to prove that f is contained in some space S such as above it suffices to assume that $f \varepsilon k(V)$. This being so, let g be a generic point over k for some component G_α of G. Then $\lambda_g f \varepsilon k(g)(V)$ is everywhere finite on V, so (again by [3, Theorem 10, p. 239]) we can write $\lambda_g f=\sum_{i=1}^N c_i(g)\phi_i$, where each $c_i(g) \varepsilon k(g)$, each $\phi_i \varepsilon k(V)$ is everywhere finite, and $c_1(g),\cdots,c_N(g)$ are linearly independent over k. Let $g_1,\cdots,g_N$ be independent generic points of G_α over k and consider the $N \times N$ matrix $(c_i(g_j))$. If we had $|c_i(g_j)|=0$, then $c_1(g_1),\cdots,c_N(g_1)$ would be linearly dependent over $k(g_2,\cdots,g_N)$, contradicting the linear disjointness of $k(g_1)$ and $k(g_2,\cdots,g_N)$ over k and the linear independence over k of $c_1(g_1),\cdots,c_N(g_1)$; thus $|c_i(g_j)| \neq 0$. Since $\lambda_{g_j} f=\sum_{i=1}^N c_i(g_j)\phi_i$, for $j=1,\cdots,N$, we can write $\phi_i=\sum_{j=1}^N d_j \lambda_{g_j} f$, where $d_1,\cdots,d_N \varepsilon k(g_1,\cdots,g_N)$. Let S be the finite dimensional vector space generated by all the ϕ_i's we obtain by letting g range over a set of generic points over k of the various components of G. If g' is generic over k for any component G_β of G, then the $g_1,\cdots,g_N$ used above could have been taken to be independent generic points over $k(g')$

of G_a, whence $\lambda_{g'}\phi_i = \sum_j d_j \lambda_{g'g_j} f \varepsilon S$, since each $g'g_j$ is generic for a component of G over k. Thus $\lambda_{g'}S \subset S$, whence $\lambda_{g'}S = S$. If γ is any point of G, we can write $\gamma = g'g''$, where g', g'' are generic points over k of components of G, so $\lambda_\gamma S = \lambda_{g'g''}S = \lambda_{g'}\lambda_{g''}S = \lambda_{g'}S = S$. Clearly $f \varepsilon S$, so S is the space we were looking for. Now let $f_1, \cdots, f_n$ be everywhere finite functions on V that are a basis for a finite dimensional invariant space S and suppose that each $f_i \varepsilon k(V)$. For any $g \varepsilon G$ we can write $\lambda_g f_i = \sum_j c_j{}^i(g) f_j$, where the $c_j{}^i(g)$'s are well-determined constants depending on g. We now show that each $c_j{}^i(g) \varepsilon k(g)$. Since $\lambda_g f_i$ is everywhere finite and defined over $k(g)$, we can write $\lambda_g f_i = \sum_{\nu=1}^N a_\nu{}^i(g)\psi_\nu$, where each $a_\nu{}^i(g) \varepsilon k(g)$, each $\psi_\nu \varepsilon k(V)$, and where $\psi_1, \cdots, \psi_N$ are linearly independent over k. Consideration of the vector space over k generated by $f_1, \cdots, f_n, \psi_1, \cdots, \psi_N$ shows that if we alter the f_i's and the ψ_ν's by suitable linear transformations with coefficients in k we can obtain $f_i = \psi_i$ for $i < s (1 \leqq s \leqq n+1)$ while $f_1, \cdots, f_n, \psi_s, \cdots, \psi_N$ are linearly independent over k. Then the equation $\sum_j c_j{}^i(g) f_j = \sum_\nu a_\nu{}^i(g)\psi_\nu$ gives

$$\sum_{j<s} (c_j{}^i(g) - a_j{}^i(g)) f_j + \sum_{j \geqq s} c_j{}^i(g) f_j = \sum_{\nu \geqq s} a_\nu{}^i(g)\psi_\nu,$$

and therefore $c_j{}^i(g) = a_j{}^i(g)$ for $j < s$, $c_j{}^i(g) = 0$ for $j \geqq s$. Thus each $c_j{}^i(g) \varepsilon k(g)$. The matrix $\Lambda_g = (c_j{}^i(g))$ is invertible (since it has the inverse $\Lambda_{g^{-1}}$) so $g \to \Lambda_g$ is a homomorphism of G into the group of $n \times n$ invertible matrices. By Theorem 3 (first applied to G_0), this is a surjective rational homomorphism from G to an algebraic group of matrices, and the homomorphism and matrix group are defined over k. In particular, the $c_j{}^i$'s are everywhere finite rational functions on the various components of G. In the case where $k(f_1, \cdots, f_n) = k(V)$, clearly the matrix group itself operates on V, so if G operates faithfully on V, then the homomorphism from G to the matrix group is a biregular isomorphism.

COROLLARY 1. *If the algebraic group G is such that G_0 is linear, then G is linear. If G is linear and defined over k, then G admits a biregular isomorphism defined over k to an algebraic group of matrices (which is therefore also defined over k).*

Let G be defined over k and let G_0 be linear. If G is finite, we may use the regular representation of G, so assume $\dim G > 0$. Let V be the direct product of the various components of G, each taken once. V is non-singular and defined over k. Left translation by elements of G defines an operation of G on V (in which any $g \varepsilon G_0$ operates on each of the direct factors of V, while if $g \varepsilon G$, $g \not\varepsilon G_0$, then g permutes the various direct factors),

and this operation is clearly defined over k, regular, and faithful. Since G_0 is linear, the coordinate functions in a matrix representation of G_0 give a set of everywhere finite rational functions on G_0 which generate the entire function field of G_0. Hence for each component G_α of G we can find a finite set of everywhere finite rational functions on G_α which generate the entire function field of G_α. Thus there exists a finite set of everywhere finite rational functions on V which generate the entire function field of V. Now apply the theorem.

COROLLARY 2. *If there exists a surjective rational homomorphism from the algebraic group G to a linear group of the same dimension, then G is linear.*

Let $\tau\colon G \to H$ be the rational homomorphism in question. We may suppose that G, and therefore also H, is connected. Letting $f_1, \cdots, f_n$ be a set of everywhere finite rational functions on H that generate the entire function field of H and letting k be a field of definition for G, H, τ, and each f_i, we get $k(H) = k(f_1, \cdots, f_n)$. Each f_i induces under τ^{-1} an everywhere finite function in $k(G)$, which we also denote f_i. Since $k(G)$ is a finite algebraic extension of $k(H)$, we can find elements $F_1, \cdots, F_s \varepsilon k(G)$ that are integrally dependent on $k[f_1, \cdots, f_n]$ such that $k(G) = k(f_1, \cdots, f_n, F_1, \cdots, F_s)$. Since each F_j is everywhere finite on G, we can apply the theorem to G operating on itself by left translation.

COROLLARY 3. *If G is an algebraic group that is defined over k, then there exists a connected normal algebraic subgroup D of G, also defined over k, such that G/D is linear and such that the kernel of any rational homomorphism from G to a linear group contains D.*

If D_1 is the kernel of a rational homomorphism from G to a linear group, Corollary 2 shows that G/D_1 is linear. If D_2 is another normal algebraic subgroup of G such that G/D_2 is linear, then the kernel of the obvious rational homomorphism from G into the linear group $(G/D_1)\times(G/D_2)$ is $D_1 \cap D_2$, so $G/(D_1 \cap D_2)$ is linear. Hence there exists a smallest normal algebraic subgroup D of G such that G/D is linear. By Corollary 2, G/D_0 is linear, so D is connected. It remains to show that D is defined over k, and for this we may suppose G connected. Let $\tau\colon G \to G/D$ be the natural homomorphism, let G and τG $(= G/D)$ operate on themselves by left translation, and let S be a finite dimensional invariant vector space of everywhere finite rational functions on τG that generates the entire function field of τG. If $g \varepsilon G$, then λ_g, $\lambda_{\tau g}$ are automorphisms of the function fields of G, τG respectively, and for any $f \varepsilon S$ we have $f\tau$ everywhere finite on G and $\lambda_g(f\tau) = (\lambda_{\tau g} f)\tau$.

Hence S_τ is a finite dimensional vector space of everywhere finite rational functions on G that is invariant under each λ_g. If Λ_g, $\Lambda_{\tau g}$ are the linear transformations induced by λ_g, $\lambda_{\tau g}$ respectively on S_τ, S, and if we choose corresponding bases for S_τ and S, then we have a matrix equality $\Lambda_g = \Lambda_{\tau g}$. If $S' \supset S_\tau$ is a finite dimensional invariant vector space of everywhere finite rational functions on G and Λ'_g is the linear transformation on S' induced by λ_g, then the map $g \rightarrow \Lambda'_g$ is a surjective rational homomorphism from G to a linear group G'. But $\tau g \rightarrow \Lambda_{\tau g}$ is a biregular isomorphism and Λ_g is merely the restriction of Λ'_g to S_τ, so we have a sequence of rational homomorphisms $G \rightarrow G' \rightarrow \tau G$. Since G' is linear, the kernel of $G \rightarrow G'$ must be D, so $G' \rightarrow \tau G$ is an isomorphism. But τ is separable, so G' is biregularly isomorphic to τG. Now note that S' could have been taken to have a basis consisting of functions in $k(G)$, in which case the map $G \rightarrow G'$ can be taken to be defined over k. This map being separable, Proposition 1, Corollary shows that its kernel D is a rational cycle over k. Hence D is defined over k.

The above proof also shows that if G is connected and $\tau\colon G \rightarrow G/D$ is taken to be defined over k then, under the obvious identification, the set of everywhere finite functions is $k(G/D)$ is precisely the set of everywhere finite functions in $k(G)$. For it is clear that any such function on G/D is one on G, so let $f \in k(G)$ be everywhere finite on G. The space S' used in the proof could have been taken so large as to include f. Letting $f_1, \cdots, f_n \in k(G)$ be a basis for S' and writing $\lambda_g f_i = \sum_j c_j{}^i(g) f_j$, the proof shows that the $c_j{}^i$'s are everywhere finite functions on G/D, and one merely has to note that $f_i(g) = (\lambda_{g^{-1}} f)(e) = \sum_j c_j{}^i(g^{-1}) f_j(e) \in k(\tau g)$.

THEOREM 13. *Let G be a connected algebraic group defined over k and let C be its center. Then C is a k-closed normal algebraic subgroup of G and G/C is linear.*

C is known to be an algebraic subgroup of G; since it is invariant with respect to all k-automorphisms of the universal domain, C is k-closed. Let $\mathfrak{o}$, a subring of the field of all rational functions on G, be the local ring of e, and let $\mathfrak{m}$ be its maximal ideal. Let $f_1, \cdots, f_n \in \mathfrak{m} \cap k(G)$ be a set of uniformizing parameters at e. For any rational function f on G and any $g \in G$ define the function $\omega_g f$ by $\omega_g f(p) = f(g^{-1}pg)$. If $g_1, g_2 \in G$, then $\omega_{g_1}\omega_{g_2} = \omega_{g_1 g_2}$, so the map $g \rightarrow \omega_g$ is a homomorphism from G to a group of automorphisms of the function field of G leaving constants fixed. For any $g \in G$, $\omega_g \mathfrak{o} = \mathfrak{o}$, so $\omega_g \mathfrak{m}^\nu = \mathfrak{m}^\nu$ for any integer $\nu > 0$. $\mathfrak{o}/\mathfrak{m}$ is the field of constants, so $\mathfrak{m}/\mathfrak{m}^\nu$ is a vector space over the constants having as a basis the various elements $\bar{f}_1{}^{i_1} \cdots \bar{f}_n{}^{i_n}$, where $i_1, \cdots, i_n$ are integers $\geqq 0$ of positive sum $< \nu$ and where

$\bar{f}$ denotes the residue class of a function $f \varepsilon \mathfrak{m}$ in the natural map $\mathfrak{m} \rightarrow \mathfrak{m}/\mathfrak{m}^\nu$. For any $g \varepsilon G$, ω_g induces an invertible linear transformation $\bar{\omega}_g$ on $\mathfrak{m}/\mathfrak{m}^\nu$. Considering the action of ω_g on $\mathfrak{o} \cap k(g)(G)$ shows that $\bar{\omega}_g \bar{f}_i$ is of the form

$$\bar{\omega}_g \bar{f}_i = \sum_{i_1 + \cdots + i_n < \nu} c^{(i)}{}_{i_1 \cdots i_n}(g) \bar{f}_1{}^{i_1} \cdot \cdot \cdot \bar{f}_n{}^{i_n},$$

where each $c^{(i)}{}_{i_1 \cdots i_n}(g) \varepsilon k(g)$. By Theorem 3, the map τ: $g \rightarrow \bar{\omega}_g$ is a rational homomorphism defined over k of G into a group of invertible square matrices, and hence τG is an algebraic group of matrices. In particular, each $c^{(i)}{}_{i_1 \cdots i_n}$ is an everywhere defined rational function on G. The kernel of τ clearly contains C. If $\gamma \varepsilon G - C$, then there exists a function $f \varepsilon \mathfrak{m}$ such that $\omega_\gamma f \neq f$, so for ν sufficiently large we have $\bar{\omega}_\gamma \bar{f} \neq \bar{f}$. Thus, for ν sufficiently large, $\bar{\omega}_g = 1$ only if g is on a closed subset of G not passing through γ. This being true for each $\gamma \varepsilon G - C$, for sufficiently large ν the kernel of τ is precisely C. In this case we have a rational isomorphism from G/C to the linear group τG. By the previous Corollary 2, G/C is linear.

COROLLARY 1. *If G is connected and D is as in the previous Corollary 3, then $D \subset C$.*

COROLLARY 2. *If the connected algebraic group G has a commutative normal algebraic subgroup H such that G/H is an abelian variety, then G is commutative.*

We have a natural surjective rational homomorphism from G/H to G/HC, so G/HC is an abelian variety. Since we also have a surjective rational homomorphism from the linear group G/C to G/HC, the latter consists of only one element. Thus $G = HC$. That is, $G = C$.

PROPOSITION 3. *Let τ be a rational map of the connected algebraic group G into the abelian variety A such that $\tau e = 0$. Then τ is a homomorphism.*

Consider the rational map of $G \times G$ into A defined by $g_1 \times g_2 \rightarrow \tau(g_1 g_2)$. By [6, Cor., p. 32] we can write $\tau(g_1 g_2) = \tau_1 g_1 + \tau_2 g_2$, where τ_1, τ_2 are rational maps of G into A. By [6, Theorem 6, p. 27], τ, τ_1, τ_2 are everywhere defined. Assuming, as we may, that $\tau_1 e = 0$, we get $\tau_2 e = \tau e = 0$. Thus $\tau_1 g_1 = \tau(g_1 e) = \tau g_1$, and similarly $\tau_2 g_2 = \tau g_2$. That is, $\tau(g_1 g_2) = \tau g_1 + \tau g_2$.

The following is a slight generalization of Weil's abstract analogue of the Poincaré theorem on complete reducibility [6, Theorem 26, p. 94], itself contained implicitly in [5, § 5].

THEOREM 14. *Let G be a connected commutative algebraic group and V a principal space with respect to G, all defined over k. Then there exists a rational map $\phi\colon V \to G$ such that ϕ is defined over k and if g, v are independent generic points of G, V respectively over k, then $\phi(gv) = \phi(v) + n \cdot g$, where n is an integer $\neq 0$.*

Here G operates regularly on V, and if k, g, v are as above and T is the locus on $V \times V$ of $v \times gv$ over k, then there is an everywhere defined rational map $\theta\colon T \to G$ such that θ is defined over k and $\theta(v \times gv) = g$. If $p \varepsilon V$ and $p' \varepsilon V$ is in the closure of the orbit Gp, then $p \times p' \varepsilon T$, so $p' = (\theta(p \times p'))p$; hence all orbits on V are closed. Thus the orbit Gv is precisely the locus on V of gv over $k(v)$. Let τ be the natural rational map from V to its variety of G-orbits, τ being taken to be defined over k. Then $k(v)$ is a regular extension of $k(\tau v)$, so let V' be the locus on V of v over $k(\tau v)$. Since $\tau v = \tau(gv)$, the natural k-isomorphism of $k(v)$ and $k(gv)$ is also a $k(\tau v)$-isomorphism, so V' is also the locus of gv over $k(\tau v)$. Hence $V' \supset Gv$. On the other hand, if v' is any generic point of V' over $k(\tau v)$, then $v' \varepsilon Gv$ (for, by a dimension argument, v' is generic for V over k and $\tau v' = \tau v$). Thus $V' = Gv$; in particular, Gv is defined over $k(\tau v)$. If $\sum_{i=1}^{n} x_i$ is a positive zero-cycle on Gv that is rational over $k(\tau v)$, then for each i we have $v \times x_i \varepsilon T$, so we can form $\sum_i \theta(v \times x_i)$, the sum being taken in G. Using the commutativity of G, the main theorem on symmetric functions [6, Theorem 1, p. 15] shows that $\phi(v) = \sum_i \theta(v \times x_i)$ is rational over $k(v)$. Since $\tau(gv) = \tau v$, we get

$$\phi(gv) = \textstyle\sum_i \theta(gv \times x_i) = \sum_i \theta(v \times x_i) - n \cdot g = \phi(v) - n \cdot g.$$

COROLLARY. *Let the abelian variety A be an algebraic subgroup of the connected algebraic group G, both defined over k. Then A is contained in the central subgroup D of G and there exists a connected k-closed algebraic subgroup G_1 of G such that $G = G_1A$ and $G_1 \cap A$ is finite. If G is non-complete, so is G_1.*

In the natural homomorphism from G to the linear group G/D, A must go into the identity, so $A \subset D$. Applying the theorem to the case where G, V are replaced by A, G (A operating on G by left translation) we get a rational map $\phi\colon G \to A$ such that ϕ is defined over k and such that for $g \varepsilon G$, $x \varepsilon A$ we have $\phi(xg) = \phi(g)x^n$, $n \neq 0$. ϕ is everywhere defined, so we may alter it by a translation on A to get $\phi(e) = e$, in which case ϕ is a homomorphism. Since $\phi(x) = x^n$, ϕ is surjective and its kernel meets A in only a finite number of points. Letting G_1 be the component of the identity of the

kernel of ϕ, we get G_1 k-closed, $G_1 \cap A$ finite, and $\dim G = \dim G_1 + \dim A$. Since the kernel of the natural homomorphism from $G_1 \times A$ to G is finite, we get $G = G_1 A$. Finally, if G_1 is complete, so is $G_1 \times A$, hence also G.

The given form of the following lemma is due to Chow.

LEMMA 1. *Let V be an abstract variety defined over k. Then there exists a projective variety V', a birational map τ from V' to V, both V' and τ being defined over k; and a k-closed subset F of V' such that τ is everywhere defined on $V' - F$ and no point of F corresponds under τ to a point of V. Considering τ as a set-theoretic map, the image of $V' - F$ is V and the inverse image of any point of V is a closed subset of V' that is disjoint from F.*

Let V be given by a coherent set of birational correspondences among $V_1 - F_1, \cdots, V_n - F_n$, where each V_i is a projective variety and F_i is a frontier on V_i, everything being defined over k. Let $P_1, \cdots, P_n$ be corresponding generic points of $V_1, \cdots, V_n$ respectively over k, let V' be the projective variety which is the locus of $P_1 \times \cdots \times P_n$ over k, and let τ be the natural birational map from V' to V. If τ_i is the birational map from V' to V_i defined by $\tau_i(P_1 \times \cdots \times P_n) = P_i$, then τ_i is everywhere defined. Let $F_i' = \tau_i^{-1}\{F_i\}$ and set $F = F_1' \cap \cdots \cap F_n'$. If $p \in F$, then $\tau_i p \in F_i$, so p corresponds to no point of V. On the other hand, if $p \in V' - F$ then for some i we have $p \in V' - F_i'$, so τ_i is defined at p and $\tau_i p \in V_i - F_i$; thus τ is defined at p. The last statement follows from the fact that each point of V corresponds to a nonempty closed subset of V'.

LEMMA 2. *Let V^n be a projective variety defined over k and let F be a nonempty k-closed proper subset of V. Then there exists a projective variety V' and a birational map $\tau\colon V' \to V$, all defined over k, such that τ is everywhere defined on V' and $\dim \tau^{-1}\{F\} = n - 1$.*

This is contained in [7]: Since it may be replaced by a derived normal model, V may be supposed relatively normal with reference to k. If $\dim F < n - 1$, let V_1 be the monoidal transform of V with respect to F and τ_1 the birational map from V_1 to V. Then τ_1 is everywhere defined and $\dim \tau_1^{-1}\{F\} > \dim F$ ([7, pp. 533, 520]). If necessary repeat this process.

LEMMA 3. *Let U^m be a variety, V^n a projective variety, $\phi\colon U \to V$ a generically surjective rational map, and X^{m-1} a simple subvariety of U, all defined over k. Then there exists a projective variety V', birationally equivalent over k to V, such that the map from V' to V is everywhere defined and*

such that if $\phi'\colon U \to V'$ is the map corresponding to ϕ and x is generic for X over k, then the locus over k of $\phi' x$ has dimension $\geqq n-1$.

The map ϕ enables us to identify $k(V)$ with a subfield of $k(U)$. Let θ be the valuation of $k(U)$ (with value group the ordinary integers) given by the order of vanishing of functions on U along X. Then the residue field with respect to θ of $k(U)$ is naturally k-isomorphic to $k(X)$, hence has transcendence degree over k equal to $m-1$. Now the transcendence degree of $k(U)$ over $k(V)$ is $m-n$, so the residue field of $k(U)$ with respect to θ has a transcendence degree over the residue field of $k(V)$ with respect to the valuation induced on it by θ which is $\leqq m-n$. Thus the residue field of $k(V)$ with respect to θ has transcendence degree $\geqq (m-1)-(m-n) = n-1$ over k. Hence there exist $f_1, \cdots, f_{n-1} \varepsilon k(V)$ which, considered as functions on U, are finite along X and induce functions on X which are algebraically independent over k; in particular, $f_1, \cdots, f_{n-1}$ are defined at x and assume at this point values $\xi_1, \cdots, \xi_{n-1}$ which are algebraically independent over k. Let V' be the graph of the rational map from V to the projective space of dimension $n-1$ determined by $(1, f_1, \cdots, f_{n-1})$. The birational map from V' to V is defined over k and everywhere defined on V'. Since V' is a subvariety of the direct product of two projective spaces, it is itself a projective variety. Since X^{m-1} is simple on U and V' is complete, $\phi' x$ is defined. The lemma now follows from the fact that $\xi_1, \cdots, \xi_{n-1} \varepsilon k(\phi' x)$.

THEOREM 15. *Let the connected algebraic group G operate regularly on the noncomplete variety V^n, all defined over the algebraically closed field k. Then there exists a normal projective variety V', defined and birationally equivalent to V over k, and a subvariety W^{n-1} of V', also defined over k, such that if we consider the operation of G on V' that is induced by its operation on V and let g, x be independent generic points of G and W respectively over k, then gx is defined and the rational map defined over k by $g \times x \to gx$ defines an operation of G on W.*

Let V', F, τ be as in Lemma 1; here F is nonempty. Note that if V'' is a projective variety and $\psi\colon V'' \to V'$ a birational map defined over k that is everywhere defined on V'', then V'', $F' = \psi^{-1}\{F\}$, and $\tau' = \tau\psi$ also satisfy the conclusions of Lemma 1. Hence, by Lemma 2, we may suppose that $\dim F = n-1$. Replacing V' by a derived normal model if necessary, we may suppose V' normal with reference to k, hence (absolutely) normal. In what follows, we use repeatedly the facts that a normal variety has no singular subvarieties of codimension one and that a rational map from a variety to a complete variety is defined along any simple subvariety of codimension one.

Let W_0 be a component of F of dimension $n-1$. The operation of G on V induces an operation of G on V', so apply Lemma 3 to the rational map $\phi\colon G\times V'\to V'$ defined by $\phi(g\times v)=gv$ and the subvariety $G\times W_0$ of $G\times V'$. We get a variety V'' and an everywhere defined rational map $\rho\colon V''\to V'$ (V'' and ρ being defined over k) such that if g, p are independent generic points of G, W_0 respectively over k, then $\rho^{-1}\phi$ is defined at $g\times p$ and $\dim_k\rho^{-1}\phi(g\times p)\geqq n-1$. Replacing V'' by a derived normal model if necessary, we may suppose V'' normal. Now the birational correspondence between V' *and* V'' is biregular between p and $\rho^{-1}p$, and that between $G\times V'$ and $G\times V''$ is biregular between $g\times p$ and $g\times\rho^{-1}p$. Replacing V' and W_0 by V'' and the locus over k of $\rho^{-1}p$, we have the following situation: V', F, τ satisfy the conclusions of Lemma 1, V' is normal, and F has a component W_0 of dimension $n-1$ such that if g, p are independent generic points of G, W_0 respectively over k, then gp (which is defined) has a locus W over k which has dimension $\geqq n-1$. If $gp\notin F$, then $\tau(gp)\in V$, so p corresponds under τ to the point $g^{-1}\tau(gp)\in V$, which is false. Hence $gp\in F$, implying $\dim W=n-1$. We claim that our present V' and W satisfy the demands of the theorem. For let g_1 be a generic point of G over $k(g,p)$. Then $g_1(gp)$ is defined, hence (by the lemma to Theorem 1) so is $(g_1g)p$ and they are equal. $(g_1g)p$ is generic for W over k; in particular the map $g_1\times gp\to g_1(gp)$ defines a rational map (defined over k) from $G\times W$ to W. Clearly $g_1(gp)$ is rational over $k(g_1,gp)$. But g_1^{-1} is generic for G over $k(g_1g,p)$, so $g_1^{-1}(g_1(gp))$ is defined and equals gp. Thus $k(g_1,g_1(gp))=k(g_1,gp)$. If now g_2 is generic for G over $k(g_1,g,p)$, then clearly $g_1(g_2(gp))=g_1g_2gp=(g_1g_2)(gp)$, completing the proof. We may add that for any $\gamma\in G$, the birational map T_γ on V' given by $T_\gamma v=\gamma v$ may be applied to a generic point of W over $k(\gamma)$ and then induces a birational map on W which is the same as that produced by γ in the operation of G on W; for if g, p are independent generic points of G, W_0 respectively over $k(\gamma)$, then $T_\gamma(gp)$ is defined and (as a specialization of the relation $T_\gamma(gv)=(\gamma g)v$) has the value $(\gamma g)p$ while $\gamma(gp)$ (defined according to the operation of G on W) equals $(\gamma g_1^{-1})(g_1(gp))$ (g_1 being generic for G over $k(\gamma,g,p)$), which equals $(\gamma g_1^{-1})((g_1g)p)=(\gamma g)p$, by Theorem 1, Lemma.

LEMMA 1. *Any noncomplete algebraic group has an algebraic subgroup of dimension >0 which is linear.*

If the noncomplete algebraic group G has dimension one, then G itself is linear, so we use induction on $\dim G$. We may assume G connected. Since G operates regularly on itself by left translation, Theorem 15 gives us

the existence of a normal variety V which is birationally equivalent to G and a subvariety W of V of dimension one less than G such that the operation of G on V induces an operation of G on W. Let W' be a variety birationally equivalent to W such that G operates regularly on W'. Fix a point $P \varepsilon W'$ and let H_P be the algebraic subgroup of G consisting of all g such that $gP = P$. The map $g \rightarrow gP$ is an everywhere defined rational map of G into W', and for any $\gamma \varepsilon G$ the points of G that map into γP are precisely γH_P. Since $\dim GP \leqq \dim W' < \dim G$, we get $\dim H_P > 0$. First assume that $H_P \neq G$. If H_P is noncomplete we can apply our induction assumption to H_P; if H_P is complete, Theorem 14, Corollary shows that G has another proper algebraic subgroup which is noncomplete, and we again use our induction assumption. We are thus reduced to the case where $H_P = G$ for all $P \varepsilon W'$, i.e. where the operation of G on W induced by its operation on V is trivial. Let k be an algebraically closed field of definition for G, V, the operation of G on V, and W. If g, P are independent generic points over k of G, W respectively, then the operation of G on V is defined at the point $g \times P \varepsilon G \times V$ and $gP = P$. Since V is normal, W is a simple subvariety of V. Thus we can find a point $p \varepsilon W$, p rational over k, such that p is simple on V and gp is defined (according to the operation of G on V) and equals p whenever g is generic for G over k. But any point of G can be expressed as the product of two points that are generic for G over k. Thus gp is defined and equals p for *all* $g \varepsilon G$. Let $\mathfrak{o}$ be the local ring of p in the function field of V (isomorphic to that of G), $\mathfrak{m}$ the maximal ideal of $\mathfrak{o}$. For any $g \varepsilon G$ the previously defined operator λ_g satisfies the relation $\lambda_g \mathfrak{o} = \mathfrak{o}$; therefore $\lambda_g \mathfrak{m}^\nu = \mathfrak{m}^\nu$ for each integer $\nu > 0$. The method of proof of Theorem 13 can now be applied to the present case. Since for each $g \varepsilon G$, $g \neq e$, one can find a function $f \varepsilon \mathfrak{m}$ such that $\lambda_g f \neq f$, this proof shows that we have a rational isomorphism from G into an algebraic group of matrices. Thus G itself is linear.

LEMMA 2. *If the connected algebraic group G possesses a normal algebraic subgroup H which is biregularly isomorphic to G_a or G_m and such that G/H is linear, then G is linear.*

Let τ be the natural map from G to G/H. H operates on G by the rule $h(g) = hg$, so by Theorem 10 there exists a rational map σ: $G/H \rightarrow G$ such that $\tau\sigma = 1$. For any $g \varepsilon G$ such that σ is defined at τg we have $\tau((\sigma\tau g)g^{-1}) = e$, so $(\sigma\tau g)g^{-1} \varepsilon H$. Let α be a coordinate function on H, got from the obvious coordinate function on G_a or G_m; then α is everywhere defined and finite on H and $\alpha(h_1) = \alpha(h_2)$ if and only if $h_1 = h_2$. The function $\alpha((\sigma\tau g)g^{-1})$ is a rational function on G that is defined and finite

Any connected algebraic group which is isogenous to such a direct product contains algebraic subgroups which are isogenous to the direct factors, proving half of the first statement. Conversely, if the connected algebraic group G contains connected algebraic subgroups L and A, respectively linear and abelian, such that $\dim G = \dim L + \dim A$, then A is contained in the center of G, so the map $\lambda \times \alpha \to \lambda\alpha$ is a rational homomorphism from $L \times A$ into G. $L \cap A$ being finite, this homomorphism has finite kernel, hence is surjective, so $G = LA$ is isogenous to $L \times A$. L is normal in G. G/L, being isogenous to $A/(A \cap L)$, is abelian, so L is the maximal connected linear algebraic subgroup of G. G/A, being isogenous to $L/(L \cap A)$, is linear, so A contains all abelian subvarieties of G. If τ is any rational homomorphism of $G = LA$, then $\tau G = (\tau L)(\tau A)$, and τL and τA are respectively linear and abelian. Finally, let H be any connected algebraic subgroup of $G = LA$. Then $H/(H \cap A)$ is isogenous to HA/A, an algebraic subgroup of the linear group G/A. Similarly $H/(H \cap L)$ is isogenous to HL/L, an algebraic subgroup of the abelian variety G/L. Thus if we set $A' = (H \cap A)_0$, $L' = (H \cap L)_0$, we get H/A' and H/L' respectively linear and abelian. The common rational homomorphic image $H/L'A'$ of both H/A' and H/L' is therefore both linear and abelian. Thus $H = L'A'$.

Note that not all such groups $G = LA$ are direct products. For example, we can find a connected commutative linear group L, an abelian variety A, and elements $\lambda \in L$, $\alpha \in A$ of finite order $n > 1$ and then let G be the factor group of $L \times A$ by the subgroup generated by $\lambda \times \alpha$. Then the images of $L \times 0$ and $e \times A$ in G meet in more than one point.

COROLLARY 7. *Notations being as in Corollary 5 and H being any normal algebraic subgroup of G, the algebraic group G/H is isogenous to the direct product of a connected linear group and an abelian variety if and only if $H \supset (L \cap D)_0$.*

First suppose that $G/H = \Lambda\Delta$, where Λ is a connected linear group and Δ is an abelian variety. Let L', D' be the inverse images of Λ, Δ respectively in the natural homomorphism from G to G/H. Then $G/L' = (G/H)/(L'/H)$ $= \Lambda\Delta/\Lambda$, which is isogenous to the abelian variety $\Delta/(\Delta \cap \Lambda)$, so G/L' is abelian, whence $L' \supset L$. Similarly, $G/D' = (G/H)/(D'/H) = \Lambda\Delta/\Delta$ is isogenous to the linear group $\Lambda/(\Lambda \cap \Delta)$, so $D' \supset D$. Hence $H \supset (L' \cap D')_0$ $\supset (L \cap D)_0$. As a result of the previous corollary, to prove the converse it suffices to take $H = L \cap D$. Then $G/H = LD/H = (L/H)(D/H)$. Here L/H is linear and $D/H = D/(L \cap D)$ is isogenous to $DL/L = G/L$ which is abelian, so D/H is an abelian variety. This ends the proof.

PROPOSITION 4. *Let G be a commutative algebraic group defined over the field k and let H be a connected k-closed algebraic subgroup of G. If the characteristic of k is $p \neq 0$ suppose also that H possesses only a finite number of elements of order p. Then H is defined over k.*

Since H is a k-closed variety, it is defined over a purely inseparable algebraic extension k' of k, so we need only consider the case where k has characteristic $p \neq 0$. Then the rational endomorphism of H given by $h \rightarrow p \cdot h$ has finite kernel, hence is surjective, so $h \rightarrow p^{\nu} \cdot h$ is surjective for any $\nu \geqq 0$. We now use an idea of Chow. If x is generic for H over k', so is $p^{\nu} \cdot x$. Fix an integer μ so large that $k(k(x))^{p^{\mu}}$ is separably generated over k and then take ν so large that the zero-cycle $p^{\nu}(x)$ on G is rational over $k(k(x))^{p^{\mu}}$. By [6, Theorem 1, p. 15] the point $p^{\nu} \cdot x$ is rational over $k(k(x))^{p^{\mu}}$, so $k(p^{\nu} \cdot x)$ is separably generated over k. Also, k' is algebraically closed in $k'(x) \supset k(p^{\nu} \cdot x)$, so k is algebraically closed in $k(p^{\nu} \cdot x)$. Therefore $k(p^{\nu} \cdot x)$ is a regular extension of k. Hence H is defined over k.

COROLLARY. *Let the connected algebraic group G be defined over k. Then $(L \cap D)_0$ and the maximal abelian subvariety of G are also defined over k.*

Each of these subgroups is left invariant by any k-automorphism of the universal domain, hence is k-closed. But each is an algebraic subgroup of the commutative group D, which is defined over k and has only a finite number of elements of any given finite order.

NORTHWESTERN UNIVERSITY.

REFERENCES.

[1] M. Rosenlicht, "Generalized jacobian varieties," *Annals of Mathematics*, vol. 59 (1954).

[2] ———, "Automorphisms of function fields," *Transactions of the American Mathematical Society*, vol. 79 (1955).

[3] A. Weil, *Foundations of Algebraic Geometry*, American Mathematical Society Colloquium Publications, New York, 1946.

[4] ———, "On algebraic groups of transformations," *American Journal of Mathematics*, vol. 77 (1955).

[5] ———, "On algebraic groups and homogeneous spaces," *American Journal of Mathematics*, vol. 77 (1955).

[6] ———, *Variétés Abéliennes et Courbes Algébriques*, Paris, 1948.

[7] O. Zariski, "Foundations of a general theory of birational correspondences," *Transactions of the American Mathematical Society*, vol. 53 (1943).

ON THE ARTIN ROOT NUMBER.*[1]

By B. DWORK.

Let X be an arbitrary character of the Galois group, $G(K/k)$, of a normal extension, K, of an algebraic number field k. The possibility of determining the arithmetic structure of the Artin root number, $W(X)$, defined by the functional equation [1] of the Artin L-series, $L(s, X, K/k)$, is suggested by the recent arithmetic characterization of $W(X)$ for X linear [2, 3, 4]. If X is linear then $W(X)$ shall be referred to as an "abelian root number."

It is the purpose of this paper to show that much of the abelian theory goes over with little modification. The main result is that if $\mathfrak{p}$ is a prime of k and if $X_\mathfrak{p}$ is the character of the local[2] Galois group, $G(Kk_\mathfrak{p}/k_\mathfrak{p})$, obtained from the restriction of X to a $\mathfrak{p}$ decomposition subgroup of $G(K/k)$ by means of the natural isomorphism between the local Galois group and the decomposition subgroup, then to within a multiplicative factor, ± 1, there exists a well defined local root number, $W(X_p)$, with factor group, linearity and induced character properties such that $W(X) = \prod W(X_\mathfrak{p})$, the product being over all primes of k. Thus (except for the question of sign to be discussed later) it is enough to determine the arithmetic structure of "irreducible local root numbers," $W(\theta)$, where θ is an irreducible character of $G(Kk_\mathfrak{p}/k_\mathfrak{p})$. If $\mathfrak{p}$ is a finite prime and $f(\theta)$ is the (local) conductor of θ, let $m(\theta) = (\mathrm{ord}_\mathfrak{p} f(\theta))/\theta(1)$. It will be shown that $W(\theta)$ is a root of unity unless $m(\theta) = 1$ in which case it is the ratio between a classical Gauss sum and its absolute value.

An immediate consequence is the integrality of Galois Gauss sums as conjectured by Hasse [5], p. 40. Hasse's conjecture (op. cit.) concerning the field in which $W(X)$ lies is treated in Theorem 7 and its corollaries.

Some remarks about the presentation are in order. While the group

* Received November 3, 1955.

[1] A summary of most of these results appears in the author's "The local structure of the Artin root number," *Proceedings of the National Academy of Science, U. S. A.*, vol. 41 (1955), pp. 754-756. Theorem 4a and its consequences form the substance of the author's dissertation, "On the root number in the functional equation of the Artin-Weil L-series, Columbia University (1954), (unpublished).

[2] "Global" and "local" are used to distinguish between algebraic number fields and their completion under a valuation (Archimedean or non-Archimedean).

theoretical discussion of Section 1 may for our immediate purpose be restricted to nilpotent groups, the more extended result has a bearing on the problem of obtaining the group theoretical properties of the Artin root number without analysis. It should also be noted that the proof of Theorem 4 below is reserved for a future paper so as to avoid excessive preoccupation at this time with purely arithmetic computations.

I am indebted to John Tate for his advice and encouragement during this investigation.

1. Group theoretical considerations. Artin's concept of extending functions defined on linear characters is easily abstracted (cf. [5], § 1).

Definition 1. A set of finite groups is said to be a *family* if it is closed under the process of forming subgroups and factor groups, it being understood that if G is a group in the set and L is a subgroup of G which contains an invariant subgroup, H, of G then L/H is identified with the image of L under the natural mapping of G into G/H, the identification to be done by means of the natural isomorphism between the two groups.

In the statement of the above definition, if L is a subgroup of G which does not contain H then $L/(H \cap L)$ is not to be identified with the image of L in G/H. In discussing families our main concern is with the characters (i.e. the traces of matrix representations with coefficients in the field of complex numbers) of the groups in the family and in particular with the mappings of the characters into some fixed abelian group. The purpose of the identification in the definition is simply to impose a natural restriction on these mappings. Of course this is achieved only if the sets of characters of identified groups are also identified in the obvious way and this further identification is to be understood. For the purpose of this section there is no need to identify the naturally isomorphic groups, G/H and $(G/N)/(H/N)$, where G is a group in the family and H is an invariant subgroup of G which contains N, an invariant subgroup of G.

Definition 2. A function, F_0, defined on the set of linear characters of the groups in a family, Δ, and taking its values in some fixed, abelian, (multiplicative) group is said to be *extendable with respect to* Δ (or Δ-extendable) if it can be extended to a function, F, on the set of all characters of groups in Δ with linearity, factor group and induced character properties. Specifically, if $G \varepsilon \Delta$ and L is a subgroup of G then

(a) If L is invariant in G and θ is a character of G/L then $F(\theta)$

$=F(\theta \circ \phi)$, $\theta \circ \phi$ being the character of G obtained by composing θ with ϕ, the natural homomorphism of G onto G/L.

(b) If X and X' are characters of G then $F(X+X')=F(X)F(X')$.

(c) If θ is a character of L and X is the character of G induced by θ then $F(\theta)=F(X)$.

It may be noted that condition (c) implies

(c′) If X is a character of L and $\sigma \varepsilon G$ then $F(X^{\sigma})=F(X)$, where X^{σ} denotes the character $x \rightarrow X(\sigma x \sigma^{-1})$ of the group $\sigma^{-1}L\sigma$.

It is an immediate consequence of Brauer's fundamental theorem on induced characters, [6], that if F exists then it is completely determined by F_0.

While the theory of non-abelian L-series, [1], depends entirely upon the concept of extendable functions, the purely group theoretical problem of characterizing such functions has received no attention. The solution of this problem for families of solvable groups or at least for groups satisfying the conditions imposed by Hilbert theory on the galois groups in local number theory would give the group theoretical structure of the Artin root number by purely arithmetic methods. Unfortunately we can at this time give the solution only for supersolvable groups (i. e. groups whose principal series have cyclic prime factor groups). Some elementary properties of solvable groups are needed before the result can be stated.

LEMMA 1. *If G is a finite solvable group, H a subgroup of prime index, p, H' the maximal subgroup of H which is invariant in G then*

(a) *There exists a unique subgroup G' of G which contains H' as a subgroup of index p.*

(b) *H/H' and G/G' are cyclic groups of equal order which divides $p-1$.*

(c) *There exists an element, x, of G such that $x \not\varepsilon H$, $x^p \varepsilon H$.*

(d) *The $(p-1)$ non-trivial linear characters of G' which are trivial on H' are permuted by the inner automorphisms of G so that the domains of transitivity contain $m=(H:H')$ elements.*

Proof. It may be assumed that H is not a normal subgroup of G, hence H is its own normalizer in G. It follows that H has exactly p conjugates, $H=H_1, H_2, \cdots, H_p$. For $x \varepsilon G$ let T_x be the permutation, $H_i \rightarrow xH_ix^{-1}$ of the conjugates of H. $x \rightarrow T_x$ is a representation of G onto a transitive permutation group on p elements. The kernel of the representation consists of all

$x \varepsilon G$ such that $xyHy^{-1}x^{-1} = yHy^{-1}$ for all $y \varepsilon G$. As H is its own normalizer, the kernel is the intersection of the conjugates of H, i.e.: H'. Statements (a) and (b) now follow from the well known properties of transitive solvable groups of permutations on a prime number of elements [7], p. 77. Let x be any element of G' not in H'. Statement (c) follows from $G' \cap H = H'$. For (d) let Y be a non-trivial linear character of G' which is trivial on H'. As it is clear that Y has at most m conjugates in G, it is enough to show them to be distinct, i.e. if $x \varepsilon G$, $Y(xyx^{-1}) = Y(y)$ for all $y \varepsilon G'$ then $x \varepsilon G'$. But this hypothesis implies that $\bar{x} = x \bmod H'$ commutes with each element of G'/H' so that G'/H' lies in the center of the group, $\bar{G}$, generated by it and $\bar{x}$, whence $\bar{G}$ is abelian as $\bar{G}/(G'/H')$ is cyclic. It follows from the previously mentioned theory of permutation groups on a prime number of elements that $\bar{G} = G'/H'$, which proves the assertion.

Definition 3. A set of groups (G, H, G', H') written in this order shall be referred to as a $(p, p-1)$ configuration if the groups are related in the manner indicated in the lemma.

Definition 4. A set of groups (G, G_1, G_2, G_0) shall be referred to as a (p, p) configuration if p is prime, G_0 is an invariant subgroup of G, G/G_0 is an abelian (p, p) group and G_1 and G_2 are distinct subgroups of G which contain G_0 as a subgroup of index p.

A family, Δ, is said to be supersolvable if every group in the family is supersolvable. The main result of this section is that the (p, p) and $(p, p-1)$ configurations are the basic units from which the relations between the characters of the groups in a supersolvable family may be determined; specifically:

THEOREM 1. *If Δ is a supersolvable family and if F is a function defined on the set of linear characters of the groups in Δ having the following properties:*

(a) *F is invariant under the transformations of the linear characters produced by inner automorphisms of the groups in Δ* (cf. Definition 2, (c′)).

(b) *If (G, G_1, G_2, H) is a (p, p) configuration of groups in Δ, H is abelian and θ_i $(i = 1, 2)$ is a linear character of G_i which induces a given irreducible character of G then $F(\theta_1) = F(\theta_2)$.*

(c) *If (G, G_0, G_1, H) is a $(p, p-1)$ configuration of groups in Δ, H is abelian, θ is a linear character of G, $\theta_0 = \theta | G_0$,[3] $\theta_1 = \theta | G_1$ and $Y_1, \cdots, Y_r$*

[3] Read: The restriction of θ to G_0.

is a minimal set of non-trivial linear characters of G_1, *trivial on* H, *whose* G *conjugates cover the set of all such characters* (cf. Lemma 1), *then* $F(\theta_0) = F(\theta) \prod_{i=1}^{r} F(\theta_1 Y_i)$.

(d) F *has the factor group property* (cf. Definition 2(a)) *for linear characters.*

Then F *is* Δ*-extendable. Regardless of* Δ *these conditions are necessary for extendability.*

This theorem is a direct consequence of Theorems 1A and 1B below. In the discussion of these theorems it is to be understood that F and Δ satisfy the conditions of Theorem 1.

THEOREM 1A. *If* X *is an irreducible character of* $G \varepsilon \Delta$ *which is induced by a linear character* θ_i *of a subgroup* G_i *of* G $(i = 1, 2)$ *then*

$$F(\theta_1) = F(\theta_2). \tag{1}$$

Proof. Let H be a maximal abelian subgroup of $G_1 \cap G_2$ which is invariant in G. The theorem is trivial if $(G:H) = 1$. The proof is by induction on the index $(G:H)$, the idea of the proof being to use the induction hypothesis to replace the given group, G, by one in which hypotheses (b) and (c) may be applied. As θ_1 and θ_2 induce the same irreducible character of G, it follows from Mackey, [8], that there exists $x \varepsilon G$ such that θ_2 and θ_1^x (exponentiation as in Definition 2(c′)) coincide on $G_2 \cap x^{-1}G_1 x$. Hence by hypothesis (a) it may be assumed that θ_1 and θ_2 coincide on $G_0 = G_1 \cap G_2 \supset H$. As θ_1 and θ_2 induce irreducible characters of $G_1 G_2$, it follows from the converse part of the previously mentioned theorem of Mackey that they induce the same character of $G_1 G_2$. Hence by the induction hypothesis it may be assumed that $G = G_1 G_2$. Let δ be the common restriction of θ_1 and θ_2 to H. δ is invariant under G_1 and G_2 and therefore under G. As G is supersolvable there exists an invariant subgroup, M, of G which contains H as a subgroup of prime index, p. Let Φ be a matrix representation of G whose character is X, then $\Phi | H = \delta I_d$, where I_d is the unit matrix of rank $d = X(1)$. As δ is invariant under G, $\Phi(H)$ lies in the center of $\Phi(G)$, hence certainly in the center of $\Phi(M)$ while $\Phi(M)/\Phi(H)$ is cyclic, whence $\Phi(M)$ is abelian. As $\Phi(G)/\Phi(M)$ is supersolvable it follows from Taketa, [9], that the character, $A \rightarrow \mathrm{Trace}\, A$, of the group $\Phi(G)$ is induced by a linear character of a subgroup of $\Phi(G)$ which contains $\Phi(M)$. It follows without difficulty that X is induced by a linear character of a subgroup of G which contains M. Without loss in generality it may be assumed that G_2

contains M. It may be further assumed that M is not a subgroup of G_1 as otherwise by hypothesis (d) and the identifications of Definition 1, θ_1 and θ_2 may be replaced by linear characters of the subgroups G_1/M^c, G_2/M^c of the factor group, G/M^c, (M^c = commutator subgroup of M) which induce the same irreducible character of the factor group, whence (1) follows from the induction hypothesis as M/M^c is an abelian subgroup of $(G_1/M^c) \cap (G_2/M^c)$ which is invariant in G/M^c and of index $(G:M) < (G:H)$. (The identifications of Definition 1 are used only in this argument. Reference is made to this argument at several points in the remainder of the proof.)

Hence $G_1 \cap M = H$ so that G_1 is a subgroup of G_1M of index p. Let X' be the character of G_1M induced by θ_1. X' is irreducible and its restriction to H is $p\delta$. As δ is certainly invariant under G_1M it follows from a previous argument that X' is induced by a linear character, θ', of a subgroup, G', of G_1M of index p which contains M. If $G \neq G_1M$ then by the induction hypothesis $F(\theta_1) = F(\theta')$ while a previous argument shows that $F(\theta') = F(\theta_2)$ as $G' \cap G_2 \supset M$.

Hence it may be assumed that $G = G_1M$, $G_2 \supset M$, $p = (G:G_1) = (G:G_2)$, $G_1 \not\supset M$ and θ_1 and θ_2 coincide on $G_0 = G_1 \cap G_2$. Clearly $G_0 \cap M = H$ so that $p = (G:G_1) \geqq (G_2:G_0) \geqq (G_0M:G_0) = (M:H) = p$. Hence $(G_1:G_0) = (G_2:G_0) = p$ and $G_2 = G_0M$. If $G_0 = H$ then $(G:H) = p^2$, whence G/H is abelian but not cyclic (as otherwise $G_1 = G_2 = G_1G_2 = G$) so that (G, G_1, G_2, H) is a (p,p) configuration so that (1) follows from hypothesis (b). Thus it may be assumed that $G_0 \neq H$ and by the argument used to justify the assumption that $G_1 \not\supset M$, it may be further assumed that no subgroup of G_0 properly containing H is invariant in G.[4]

Summarizing: It is enough to prove (1) for the case in which $G = G_1G_2$, $p = (G:G_1) = (G:G_2) = (G_1:G_0) = (G_2:G_0)$, $G_0 = G_1 \cap G_2 \supset H$ (strict inclusion), H an abelian invariant subgroup of G, $\theta_1 | G_0 = \theta_2 | G_0$, $p^2 | (G:H)$, no subgroup of G_0 properly containing H is invariant in G.

As no further use shall be made of the group M, there will be no need in the remainder of the proof to repeat for G_1 arguments applicable to G_2 and conversely. To complete the proof it is necessary to examine the structure of G more closely. We assert: If U is a subgroup of G_2 which is invariant in G then $U' = U \cap G_0$ is also invariant in G. To prove this let $s_1, \cdots, s_p$ be a set of representatives of the right cosets of G_0 in G_2 and let $t_1, \cdots, t_p$ be a set of representatives of the right cosets of G_0 in G_1. If $u \in U'$ then $t_iut_i^{-1} \in G_0$, whence $X(u) = \sum_{i=1}^{p} \theta_2(t_iut_i^{-1}) = p\theta_1(u)$. Hence $p\theta_1(u) = \sum_{i=1}^{p} \theta_1(s_ius_i^{-1})$.

[4] This completes the proof if G is nilpotent.

If $s_i u s_i^{-1} \varepsilon G_1$ then $\theta_1(s_i u s_i^{-1}) = \theta_1(u)$, whence $s_i u s_i^{-1} \varepsilon G_1$ for $i = 1, \cdots, p$ and therefore $s_i U' s_i^{-1} \subset G_1$. It follows that $s_i U' s_i^{-1} = U'$ for each i and as U' is invariant in G_1 the assertion follows.

Let Z_i $(i = 1, 2)$ be the maximal subgroup of G_i which is invariant in G. It follows from the above assertion that $Z_i \cap G_0$ is invariant in G and contains H and therefore is H. From Lemma 1, $(G:Z_i)$ divides $p(p-1)$ but p^2 divides $(G:H)$, hence p divides $(Z_i:H)$. As $(G_1Z_2:Z_2) = (G_1:H)$ it follows that $(G:H) \geqq (G_1Z_2:H) = (G_1Z_2:Z_2)(Z_2:H) = (G_1:H)(Z_2:H)$ $\geqq p(G_1:H) = (G:H)$. Hence $G_1Z_2 = G$ and $(Z_2:H) = p$. Likewise $(Z_1:H) = p$ and $G_2Z_1 = G$. It now follows that (G_1, G_0, Z_1, H) is a $(p, p-1)$ configuration, for if M is a subgroup of G_0 which is invariant in G_1 then MZ_2 is invariant under conjugation by elements of G_1 and by elements of Z_2 and therefore by elements of $G_1Z_2 = G$, whence $MZ_2 = Z_2$ so that $M \subset G_0 \cap Z_2 = H$ and therefore H is the maximal such subgroup of G_0. Likewise (G_2, G_0, Z_2, H) is a $(p, p-1)$ configuration. Furthermore $H \subset Z_1 \cap Z_2 \subset Z_1 \cap G_0 = H$ so that letting $Z = Z_1Z_2$, we have $(Z:H) = p^2$ and therefore (Z, Z_1, Z_2, H) is a (p, p) configuration.

Let $\phi_1 = \theta_1 | Z_1$ and let s be an element of Z_2 which generates the factor group Z_2/H. Set $L = G_1 \cap s^{-1}G_1 s$, then L and sLs^{-1} are subgroups of G_1 which contain Z_1 and are of the same index in G_1. As G_1/Z_1 is cyclic it follows that they are equal, whence L is an invariant subgroup of G, i.e. $L = Z_1$. As θ_1 induces an irreducible character of G it follows from Mackey (op. cit.) that θ_1 and θ_1^s do not coincide on Z_1 but do coincide on H as δ is invariant under G. As ϕ_1 is invariant under G_1 it follows that ϕ_1 has exactly p conjugates, $\phi_1 Y_0, \phi_1 Y_1, \cdots, \phi_1 Y_{p-1}$, where $Y_0, \cdots, Y_{p-1}$ are the linear characters of Z_1 which are trivial on H. A similar statement holds for $\phi_2 = \theta_2 | Z_2$. It is easily verified that ϕ_1 induces the character $X | Z$ of Z. $X | Z$ is irreducible as now follows from the converse part of the last mentioned theorem of Mackey (it is for this that s is chosen in Z_2). It follows from hypothesis (b) that

$$F(\phi_1) = F(\phi_2). \tag{2}$$

Furthermore, letting $\theta_0 = \theta_1 | G_0$, it follows from hypothesis (c) that

$$F(\theta_0) = F(\theta_1) \prod_{i=1}^{r} F(\phi_1 Y_i) \tag{3}$$

for a suitable indexing of the Y_i. From Lemma 1, $r = (p-1)/(G_0:H)$, but the $\phi_1 Y_i$ are conjugates of ϕ_1, whence by hypothesis (a),

$$F(\theta_0) = F(\theta_1)[F(\phi_1)]^r. \tag{4}$$

Clearly (4) remains valid if the subscript 1 is replaced by 2, hence (1) follows from (2).

This completes the proof of Theorem 1A. F may be extended in a natural way to all irreducible characters of groups in Δ as if X is such a character then it is induced by a linear character θ of a subgroup and while the choice of θ need not be unique, $F(\theta)$ depends only upon X. The extension to composite characters so that the linearity condition is satisfied is obvious. The same symbol, F, will be used to denote the extended function. The factor group property follows easily from hypothesis (d). Some preliminary results are needed to verify the induced character property (and so complete the proof of Theorem 1). In these elementary lemmas the groups referred to are understood to be arbitrary finite groups.

LEMMA 2. *Let X be an irreducible character of a group G whose restriction to an invariant subgroup, H contains a linear character, δ. If G' is the subgroup of G which leaves δ invariant then $X|G'$ contains just one irreducible character, Θ, whose restriction to H contains δ. Furthermore X is induced by Θ.*

Proof. Certainly $X|G'$ contains an irreducible character, Θ, of G' whose restriction to H contains δ. Clearly $\Theta|H = \Theta(1)\delta$. If $t \in G$, $t \notin G'$ then $\delta^t \neq \delta$ so that $\Theta^t|H$ and $\Theta|H$ have in common no irreducible character of H, hence certainly the restrictions of Θ^t and Θ to $G' \cap t^{-1}G't$ have no irreducible character of that group in common. It follows from the previously mentioned results of Mackey that Θ induces X. Hence $X|H$ contains δ exactly $\Theta(1)$ times so that no other irreducible character of G' which occurs in $X|G'$ can have δ in its restriction to H.

LEMMA 3. *Let X be the character of a group G which is induced by a linear character, θ, of a subgroup L. If δ is the restriction of θ to an invariant subgroup, H, of G contained by L and if G' is the subgroup of G which leaves δ invariant then the characters, Θ of G' and X of G, induced by θ, have decompositions, $\Theta = \sum_{i=1}^{r} a_i\Theta_i$, $(a_i \neq 0)$ $X = \sum_{i=1}^{r} a_iX_i$, X_i is induced by Θ_i, into distinct irreducible characters of their respective groups.*

Proof. Clearly L is a subgroup of G'. Let $\Theta = \sum_{i=1}^{r} a_i\Theta_i$ be the decomposition of Θ into distinct irreducible characters of G': As Θ induces X, each irreducible character, X_j, of G occurring in X lies in the character of G induced by one of the Θ_i. If Θ_i and X_j are so related then by Frobenius,

Θ_i occurs in $X_j|G'$, whence by Lemma 2, Θ_i is the only irreducible character of G' which can be so related and furthermore Θ_i induces X_j. The lemma follows directly.

The induced character property may now be demonstrated.

THEOREM 1B. *Let θ be a character of a subgroup, L, of a group $G \varepsilon \Delta$. Let X be the character of G induced by θ, then $F(X) = F(\theta)$.*

Proof. It may be assumed that θ is irreducible and therefore it may be assumed that θ is linear. Let H be a maximal abelian subgroup of L which is invariant in G. The theorem is trivial if $(G:H) = 1$ and the proof is by induction on this index. Once again the idea of the proof is to use the induction hypothesis to replace G by a group in which hypothesis (c) may be used. Let $\delta = \theta|H$ and let G' be the subgroup of G which leaves δ invariant. Using the notation and result of Lemma 3 it is clear that $F(X) = F(\Theta)$, whence by the induction hypothesis it may be assumed that $G' = G$, i.e. δ is invariant under G. As before it follows from the induction hypothesis and the factor group property that it may be assumed that H is the maximal invariant subgroup of G in L. As before there exists an invariant subgroup, M, of G which contains H as a subgroup of prime index, p. From the preceding remark $L \not\supset M$, whence $L \cap M = H$. Let $LM = N$, then $(N:L) = p$. Let X' be the character of N which is induced by θ. If $N \neq G$ then by the induction hypothesis $F(\theta) = F(X')$. As $X'|H = X'(1)\delta$ it follows from a previous argument that each irreducible character of N in X' is induced by a linear character of a subgroup of N which contains M. It follows from the linearity and factor group properties of F and the induction hypothesis that $F(X') = F(X)$. Hence it may be assumed that $G = ML$, $(G:L) = p$. As H is the maximal invariant subgroup of G in L it follows that (G, L, M, H) is a $(p, p-1)$ configuration. Let $m = (L:H) = (G:M)$ then $m|(p-1)$. We assert that there exists a linear character X_0 of G whose restriction to H is δ. This is clear if $m = 1$ as then G/H is cyclic, whence from the invariance of δ under G it follows that δ is trivial on the commutator subgroup of G. On the other hand if $m > 1$ then $sLs^{-1} \subset L$ for no element $s \varepsilon M$, $s \not\varepsilon H$. Let t be an element of L whose coset mod H generates the cyclic group, L/H. Then $X(t) = \theta(t) \neq 0$. But each irreducible character, X_i, of G occurring in X is induced by a linear character of a subgroup, G_i, which contains M. As G/M is cyclic, G_i is invariant in G, whence $X_i(t) = 0$ unless $t \varepsilon G_i$. Hence there exists i such that $t \varepsilon G_i$ so that $G_i = G$ and therefore $X_i(1) = 1$. Having shown that X_0 exists it follows from the reciprocity law that $\theta = X_0|L$. Let $\theta_0 = X_0|M$, then θ_0 is invariant

under G. As $X|M$ is the character of M induced by δ, it follows easily that $X|M = \theta_0 + \theta_0 Y_1 + \cdots + \theta_0 Y_{p-1}$, where $Y_1, \cdots, Y_{p-1}$ are the nontrivial linear characters of M which are trivial on H. By Lemma 1, Y_i has m distinct conjugates in G, whence the same holds for $\theta_0 Y_i$ and therefore by Mackey, $\theta_0 Y_i$ induces an irreducible character of G of degree m. Each of the $r = (p-1)/m$ distinct irreducible characters of G of this type must occur at least once in X and therefore $X = X_0 + \sum_{i=1}^{r} X_i$, X_i induced by $\theta_0 Y_i$; $Y_1, \cdots, Y_r$ chosen so that their G conjugates cover the full set $Y_1, \cdots, Y_{p-1}$. By definition $F(X) = F(X_0) \prod_{i=1}^{r} F(\theta_0 Y_i)$, but the right side is $F(\theta)$ (hypothesis (c)). This completes the proof of Theorem 1B and therefore of Theorem 1, the necessity of the conditions being now clear.

The solvable group of lowest order which is not supersolvable is the tetrahedron group of order 12. Supersolvable groups are characterized by the property: Every maximal subgroup is of prime index [13]. For this reason Lemma 1 is adequate only for the study of supersolvable families. The extension of the theorem to solvable groups requires an examination of the situation discussed in Lemma 1 with H a subgroup of G of index p^r (p prime). This can be done if enough is known about solvable groups of permutations of p^r elements.

2. Field theoretic considerations. Theorem 1 permits the construction of local root numbers of characters of supersolvable local Galois groups. To remove the restriction of supersolvability without a deeper group theoretical investigation it is necessary to make use of the analytically derived group theoretical structure of the (global) root number as explained by Theorem 3 below. This is done by means of the relations between local and global number fields, specifically the existence of global number fields whose local completions may to some extent be preassigned.

THEOREM 2. *Let A be a finite normal non-cyclic extension of a local number field, B. If m (>1) is the degree over B of an intermediate field then there exists an algebraic number field, k, with normal overfield, K, such that the set, S, of all primes of k which have just one prime divisor in K has either one or m elements and in either case B is (topologically[5]) isomorphic to the completion of k at each prime in S and A is isomorphic to the completion of K at each prime of K which divides a prime in S. Furthermore the fields K, k may be so chosen that for each prime, $\mathfrak{P}$, of K*

[5] The topology is given by the valuation.

which divides a prime in S there exists a $\mathfrak{P}$-topological isomorphism $\Phi_{\mathfrak{P}}$ of K into A which maps k into B such that if $\Sigma \to \Sigma_{\mathfrak{P}}$ is the isomorphism of $G(K/k)$ onto $G(A/B)$ defined by

$$(\Sigma_{\mathfrak{P}} \circ \Phi_{\mathfrak{P}})(x) = (\Phi_{\mathfrak{P}} \circ \Sigma)(x) \text{ for all } x \varepsilon K$$

then the mapping $\Sigma \to \Sigma_{\mathfrak{P}}$ remains unchanged as $\mathfrak{P}$ runs through the primes of K which divide a prime in S.

Proof. Let n be the degree of A over B. Certainly there exists an algebraic number field, k, with normal extension, K, of degree n such that B is isomorphic to the completion of k at some prime, $\mathfrak{p}_0$, of k and such that A is isomorphic to the completion of K at the prime of that field which divides $\mathfrak{p}_0$. Let S_0 be the set of all primes of k other than $\mathfrak{p}_0$ which have just one prime divisor in K. S_0 is finite as K is a non-cyclic extension of k. It may be assumed that S_0 is not empty. By hypothesis there exists a subfield, L, of K such that $m = \text{degree } L/k > 1$. Let v be an integer of L which generates L over k. Let f be the irreducible monic polynomial with coefficients in k which is satisfied by v. Let $a_1, \cdots, a_m$ be any set of distinct integers in k and let $\prod_{i=1}^{m} (x - a_i) = f_0(x)$. By the approximation theorem [10, page 8] there exists a polynomial, g, of degree m with coefficients in k which is so close to f at the primes of S_0 and to f_0 at $\mathfrak{p}_0$ that

(a) g splits in $k_{\mathfrak{p}_0}$, the $\mathfrak{p}_0$ completion of k.

(b) If $\mathfrak{p} \varepsilon S_0$ and if $v_1, \cdots, v_m$ are the roots of f and $b_1, \cdots, b_m$ are the roots of g in a field containing $k_{\mathfrak{p}}$, then for suitable choice of indices, $k_{\mathfrak{p}}(v_i) = k_{\mathfrak{p}}(b_i)$, $1 \leqq i \leqq m$.

Let w be a root of the polynomial, g, in an extension field of K. Let $F = k(w)$ and let $E = KF$. We assert that the fields F, E have the required properties. For $\mathfrak{p} \varepsilon S_0$, f is irreducible in $k_{\mathfrak{p}}$ and therefore by (b), g is irreducible in $k_{\mathfrak{p}}$, hence g is irreducible in k. Thus degree $F/k = m$ and $\mathfrak{p}$ has just one prime divisor in F. By (a), $k_{\mathfrak{p}_0} \supset F$ so that $\mathfrak{p}_0$ has m prime divisors, $\mathfrak{q}, \cdots, \mathfrak{q}_m$ in F and for each of them $F_{\mathfrak{q}_i} = k_{\mathfrak{p}_0} \cong B$ while $EF_{\mathfrak{q}_i} = Kk_{\mathfrak{p}_0} \cong A$, $(1 \leqq i \leqq m)$. Hence

$$\text{degree } E/F \geqq \text{degree } A/B = n = \text{degree } K/k \geqq \text{degree } E/F.$$

It follows that $n = \text{degree } E/F$ and that each of these primes of F has just one prime divisor in E. Thus it suffices to show that the remaining primes of F have more than one prime divisor in E. If $\mathfrak{q}$ is one of these remaining

primes of F then $\mathfrak{q}$ divides a prime, $\mathfrak{p}$, of k which either lies in S_0 or has more than one prime divisor in K.

Case 1. $\mathfrak{p} \varepsilon S_0$. As previously noted $\mathfrak{q}$ is the only prime divisor of $\mathfrak{p}$ in F. Since f splits in K and therefore in $Kk_{\mathfrak{p}}$, it follows that g splits in $Kk_{\mathfrak{p}}$ so that $Kk_{\mathfrak{p}} \supset F_{\mathfrak{q}}$. Hence $EF_{\mathfrak{q}} = KF_{\mathfrak{q}} = Kk_{\mathfrak{p}}$ is of degree n over $k_{\mathfrak{p}}$ and therefore of degree n/m over $F_{\mathfrak{q}}$ which shows that $\mathfrak{q}$ has m prime divisors in E.

Case 2. $\mathfrak{p} \not\varepsilon S_0$. As $\mathfrak{p}$ has more than one prime divisor in K,

$$n > \text{degree}\, Kk_{\mathfrak{p}}/k_{\mathfrak{p}} \geqq \text{degree}\, EF_{\mathfrak{q}}/F_{\mathfrak{q}}$$

(as $EF_{\mathfrak{q}}$ is the composition of $Kk_{\mathfrak{p}}$ with $F_{\mathfrak{q}}$), which shows that $\mathfrak{q}$ has more than one prime divisor in E.

The last assertion of the theorem follows from the fact that for $1 \leqq i \leqq m$, the prime $\mathfrak{Q}_i$ of E which divides the prime q_i of F also divides that prime, $\mathfrak{P}_0$, of K which divides the prime, $\mathfrak{p}_0$, of k. Furthermore the local degree at $\mathfrak{Q}_i$ of E over K is 1. Let ϕ be a $\mathfrak{P}_0$-topological isomorphism of K into A which maps k into B and let $\Sigma \to \bar{\Sigma}$ be the isomorphism of $G(E/F)$ onto $G(A/B)$ defined by

$$(\bar{\Sigma} \circ \phi)(x) = (\phi \circ \Sigma)(x) \text{ for all } x \varepsilon K.$$

It is clear that ϕ can be extended to Φ_i, a $\mathfrak{Q}_i$-topological isomorphism of E into A. Φ_i maps F into B and the mapping $\Sigma \to \Sigma_i$ of $G(E/F)$ onto $G(A/B)$ defined by means of Φ_i is the same as the mapping $\Sigma \to \bar{\Sigma}$. This completes the proof of the theorem.

The basic problem of this section is that of characterizing functions defined on linear characters of local Galois groups which can be extended with respect to the family of all local Galois groups. Theorem 3 below (together with Theorem 1) gives a partial solution of this problem which is adequate for the theory of root numbers. Before stating the result we pause to explain the terminology.

Let Q be the field of rational numbers and for each prime p of Q (including the infinite one) let Q_p be a p-completion of Q. Let $\mathfrak{T}_p$ (resp.: $\mathfrak{T}$) be the set of all overfields of Q_p (resp.: Q) of finite degree which lie in a fixed algebraic closure of Q_p (resp.: Q), the construction being so performed that no two of these algebraic closures have any element in common. Let $\bar{\mathfrak{T}}$ be the set theoretic union of the disjoint sets, $\mathfrak{T}_p$, where p ranges over all primes of Q. The elements of $\bar{\mathfrak{T}}$ are understood to be topological fields and isomorphisms between two elements will be understood to be topological. $\bar{\mathfrak{T}}$ (resp.: $\mathfrak{T}$) shall be referred to as the *set of all local* (resp.: *global*) *number*

fields. For A, B in $\bar{\mathfrak{T}}$ (resp.: $\mathfrak{T}$), A a normal overfield of B, the Galois group, $G(A/B)$, of A over B is regarded not as an abstract group but rather as being inextricably connected with the pair (A, B) of fields. The usual Galois theoretic identifications concerning subgroups and factor groups being understood, it is clear that the set, $\bar{\mathfrak{G}}$ (resp.: $\mathfrak{G}$), of all such Galois groups is a family in the sense of Definition 1 and shall be referred to as the *family of all local* (resp.: *global*) *Galois groups.* (Precisely as in the purely group theoretical case, $G(CD/D)$ is not to be identified with $G(C/C \cap D)$ as elements of $\bar{\mathfrak{G}}$ (resp.: $\mathfrak{G}$) even when these groups are isomorphic (unless $C \supset D$).)

The set, $\bar{\mathfrak{X}}$ (resp.: $\mathfrak{X}$) of all characters of groups in $\bar{\mathfrak{G}}$ (resp.: $\mathfrak{G}$) shall be referred to as the *set of all characters of local* (resp.: *global*) *Galois groups.* Two characters in $\bar{\mathfrak{X}}$ (resp.: $\mathfrak{X}$) are said to be *equivalent* if one is obtained from the other in the natural way from a field isomorphism. Let Γ be a multiplicative abelian group, fixed throughout the discussion. A mapping, F, of $\bar{\mathfrak{X}}$ (resp.: $\mathfrak{X}$) into Γ which is constant on each class of equivalent characters is said to be a *function defined on all characters of local* (resp.: *global*) *Galois groups.* A similar interpretation is to be given to the *set of all linear characters of local* (resp.: *global*) *Galois groups* and to the concept of a *function defined on all linear characters of local* (resp.: *global*) *Galois groups.* A function defined on all linear characters of local (resp.: global) Galois groups will be said to be extendable with respect to a given family of local (resp.: global) Galois groups if the restriction of the function to the linear characters of the groups in the family is extendable with respect to the family.

If A is a local number field then a homomorphism of A^* (the multiplicative group of non-zero elements of A) into the unimodular complex numbers wtih kernel of finite index in A^* will be said to be a *multiplicative character of A* (or a *character of A^**). The set of all such homomorphisms as A ranges over $\bar{\mathfrak{T}}$ will be called the *set of all multiplicative characters of local number fields.* Two such characters are said to be *equivalent* if one is transformed into the other by a field isomorphism. A mapping, F, of all multiplicative characters of local number fields into Γ which is constant on each class of equivalent characters is said to be a *function defined on all multiplicative characters of local number fields.* It follows from local class field theory [6] that such functions may be identified with functions defined on all linear

[6] To avoid confusion with the classical norm rest symbol, it is to be understood that a prime element of $k_{\mathfrak{p}}$ (if $\mathfrak{p}$ is a finite prime) is to be associated with automorphisms of abelian over-fields which in the unramified case is simply the Frobenius substitution.

characters of local Galois groups which have the factor group property. This identification is to be understood in the following.

Finally, if K and k are elements of $\mathfrak{T}$, K being normal over k, and if $\mathfrak{p}$ is a prime of k then $G(Kk_{\mathfrak{p}}/k_{\mathfrak{p}})$ will be used to denote an element, $G(A/B)$, of $\bar{\mathfrak{G}}$ such that there exists a $\mathfrak{P}$-topological isomorphism, ϕ, of K onto a dense subset of A which maps k onto a dense subset of B, where $\mathfrak{P}$ is a prime of K which divides $\mathfrak{p}$. If X is a character of $G(K/k)$ then $X_{\mathfrak{p}}$ is to denote the character of $G(A/B)$ which is obtained by transforming (by means of ϕ) the restriction of X to a $\mathfrak{P}$-decomposition subgroup of $G(K/k)$. Of course $X_{\mathfrak{p}}$ is not a uniquely defined element of $\bar{\mathfrak{X}}$ but if F is a function defined on all characters of local Galois groups then $F(X_{\mathfrak{p}})$ is well defined.

THEOREM 3. *Let F be a function defined on all linear characters of local Galois groups with the properties:*

(a) *F is extendable with respect to every family of nilpotent local Galois groups.*

(b) *If X is a linear character of a global Galois group, $G(K/k)$, where k is an algebraic number field and K is a (finite) normal overfield, then $F(X_{\mathfrak{p}})=1$ for almost all primes, $\mathfrak{p}$, of k ($X_{\mathfrak{p}}$ being defined as indicated above.*

(c) *If in the notation of* (b), *M denotes the function $X \to \prod_{p} F(X_{\mathfrak{p}})$ defined on the set of linear characters of global Galois groups, the product being over all primes of k, then M is extendable with respect to every family of global Galois groups.*

Then it may be concluded that F is extendable with respect to the family of all local Galois groups.

Proof. It is enough to prove:

If B is a $\mathfrak{p}$-adic number field, A a normal extension of finite degree, and Δ' is the family of groups generated by the Galois group, $G(A/B)$, (i.e. the family of all Galois groups, $G(A'/B')$, where $A \supset A' \supset B' \supset B$, A' normal over B') then F is extendable with respect to Δ'.

The proof is by induction on degree A/B. It may be assumed that the Galois group, $G(A/B)$, is neither cyclic nor a p-group. Let m (>1) be the degree over B of some intermediate field. Let k and K be the algebraic number fields whose existence has been demonstrated in the last theorem, and let S be the set of all primes of k which have just one prime divisor in K. The number, r, of primes in S is either 1 or m. By the induction

hypothesis, F is extendable with respect to the family of groups generated by the local Galois group, $G(Kk_{\mathfrak{p}}/k_{\mathfrak{p}})$, for each prime $\mathfrak{p}$ of k which is not in S. Let F' denote this extended function for all $\mathfrak{p} \notin S$. Let Δ be the family of groups generated by $G(K/k)$. If $G_1 \in \Delta$ then G_1 is the Galois group, $G(U/V)$, of U over V, U and V being fields lying between k and K, U being normal over V. If X is a character of G_1 then for each prime, $\mathfrak{p}$, of k not in S we set

$$H_{\mathfrak{p}}(X) = \prod_{\mathfrak{q}|\mathfrak{p}} F'(X_{\mathfrak{q}}), \tag{4}$$

the product being over all primes, $\mathfrak{q}$, of V which divide $\mathfrak{p}$ and $X_{\mathfrak{q}}$ being related to X in the usual way. For each prime, $\mathfrak{p}$, of k not in S, (4) defines a function on the characters of the groups in Δ. As F' satisfies the conditions of Definition 2 with respect to those local groups of which the $X_{\mathfrak{q}}$ in (4) are characters, it follows from an argument of Artin [11, pages 3-5] that $H_{\mathfrak{p}}$ satisfies the conditions of Definition 2 with respect to the family Δ. If X is a linear character of G_1 then by hypothesis (b),

$$H_{\mathfrak{p}}(X) = 1 \text{ for almost all primes, } \mathfrak{p}, \text{ of } k. \tag{5}$$

By Brauer (loc. cit.), (5) holds for all characters of G_1. Hence if we set

$$H(X) = \prod_{\mathfrak{p} \notin S} H_{\mathfrak{p}}(X), \tag{6}$$

the right side being a product over all primes of k not in S, then H is a well defined function of the characters of the groups in Δ which satisfies all the conditions of Definition 2, and therefore (using M to denote the extension of the function, M, (hypothesis (c)))

$$N(X) = M(X)/H(X) \tag{7}$$

defines a function, N, on these characters which also satisfies these conditions, i. e. the restriction of N to the linear characters of the groups in Δ is Δ-extendable.

If X is a linear character of G_1 then

$$N(X) = \prod_{\mathfrak{p} \in S} \prod_{\mathfrak{q}|\mathfrak{p}} F(X_{\mathfrak{q}}), \tag{8}$$

the combined product being over all primes of V which divide a prime in S. Let $\bar{S}$ be the set of primes of K which divide a prime in S. For each $\mathfrak{P} \in \bar{S}$ let $\Phi_{\mathfrak{P}}$, a $\mathfrak{P}$-topological isomorphism of K into A, be chosen as in the last assertion of Theorem 2. For each $\Phi_{\mathfrak{P}}$ there exists a one to one correspondence between groups in Δ and groups in Δ' and by the choice of the $\Phi_{\mathfrak{P}}$ this correspondence does not change as $\mathfrak{P}$ runs through $\bar{S}$. Likewise there exists a one to one correspondence $X \leftrightarrow X'$ between the characters of the groups in Δ

and the characters of the groups in Δ' which can be described in terms of $\Phi_{\mathfrak{P}}$ but does not depend upon the choice of $\mathfrak{P}$ in $\bar{S}$. It follows from the topological nature of this construction that if X is the linear character of G_1 in (8) then $F(X') = F(X_{\mathfrak{q}})$ for each prime, $\mathfrak{q}$, of V which divides a prime in S. As each prime in S has just one prime-divisor in V, it follows that

$$N(X) = (F(X'))^r \text{ for } X \text{ linear.} \tag{9}$$

It is now clear that the function $F^r : X' \to (F(X'))^r$ on the linear characters of the groups in Δ' is Δ'-extendable. It may be assumed that it is impossible to construct the fields, K, k, so that $r = 1$, hence it may be assumed that $r = m$, the degree over B of a field (not B) lying between A and B. As $G(A/B)$ is not a p-group, the indices of the Sylow subgroups from a set of relatively prime integers each greater than one, which proves the theorem.

While the theorem is adequate for our purposes it is clear that condition (a) is stronger than necessary.

COROLLARY. *The theorem remains valid if condition* (a) *is replaced by either:*

(a′) *F is extendable with respect to every family of local Galois groups which consists of cyclic groups and p-groups,*

or

(a″) *F is extendable with respect to every family of local Galois groups which consists of p-groups and of the Galois groups of unramified extensions.*

The proof is clear but it should be noted that the last sentence of the proof of the theorem shows why p-groups require special treatment. The special treatment of cyclic groups (or of the Galois groups of unramified extensions) is necessitated by the requirement of finiteness of S_0 in the proof of Theorem 2. It should be observed that (a″) corresponds to an obvious modification of that theorem.

3. Arithmetic considerations. The local abelian root numbers must now be defined.

Definition 5. If θ is a multiplicative character of a local number field, B, then the local abelian root number, $R(\theta)$, is defined by

$$\begin{aligned} R(\theta) &= 1 \quad \text{if } B \text{ is complex,} \\ &= 1 \quad \text{if } B \text{ is real and } \theta(-1) = 1, \\ &= -i \quad \text{if } B \text{ is real and } \theta(-1) = -1, \\ &= (Nf(\theta))^{-\frac{1}{2}} \sum_{x \in U/(1+f(\theta))} \bar{\theta}(x/A_\theta)\phi(x/A_\theta) \quad \text{if } B \text{ is } \mathfrak{p}\text{-adic,} \end{aligned}$$

where (in explanation of the $\mathfrak{p}$-adic case)

$f(\theta) =$ conductor of θ,

$Nf(\theta) =$ absolute norm of $f(\theta)$,

A_θ is an integer of B which generates the ideal, $f(\theta)\mathfrak{D}_B$, $\mathfrak{D}_B$ being the absolute different of B,

ϕ is the "standard" additive character of $B: x \rightarrow \exp(2\pi i Y(S(x)))$, S being the absolute trace mapping B onto Q_p, the corresponding completion of the field of rational numbers, Y being a mapping of Q_p into the rationals such that for each $t \in Q_p$, $Y(t) - t$ is a p-adic integer and $Y(t)$ is a rational number whose denominator is a power of p,

U is the group of units in B,

and the sum is over a set of representatives of the cosets in U of the subgroup of all units congruent to 1 modulo $f(\theta)$, it being understood that if θ is unramified then $1 + f(\theta) = U$.

It follows from the arithmetic characterization of abelian root numbers, [4], that $R(\theta)$ is a unimodular complex number in any case and is even a root of unity unless B is $\mathfrak{p}$-adic and $f(\theta) = \mathfrak{p}$, in which case $R(\theta)$ is the ratio between a classical Gauss sum and its absolute value.

To simplify the statement of the arithmetic results: If A is a finite overfield of the local number field, B, then a character, θ, of B^* (i.e. a multiplicative character of B) is said to divide a character, Θ, of A^* (written $\theta|\Theta$) if Θ is obtained from θ by composition with the relative norm, $N_{A/B}$.

THEOREM 4. *Let B be a local number field,*

(a) *If A is a cyclic overfield of prime degree and Θ is a character of A^* divisible by a character of B^* then*

$$[\prod_{\theta|\Theta} R(\theta)]/R(\Theta) = V(A/B) \tag{10}$$

$$[\prod_{\theta|\Theta} \theta(-1)]/\Theta(-1) = [V(A/B)]^2 \tag{11}$$

the products being over all characters of B^* which divide Θ, $V(A/B)$ being a fourth root of unity which is independent of Θ. Furthermore $V(A/B) = 1$ if degree A/B is odd.

(b) *If A is an abelian overfield of B with Galois group of type (p, p) and if θ_j $(j = 1, 2)$ is a character of A_j^*, A_j being an intermediate field of*

degree p over B, such that θ_1 and θ_2 divide the same character of A^*, neither being divisible by any character of B^* then

$$R(\theta_1)V(A_1/B) = R(\theta_2)V(A_2/B) \tag{12}$$

$$\theta_1(-1)[R(\theta_1)]^2 = \theta_2(-1)[R(\theta_2)]^2. \tag{13}$$

COROLLARY. *Equations* (10) *and* (11) *are also valid if A is any abelian extension of B.*[7] *If C is an immediate field then*

$$V(A/B) = V(A/C)[V(C/B)]^{\text{degree } A/C}. \tag{14}$$

The corollary is derived from the theorem by a simple induction argument. The theorem is proven by the arithmetic characterization of local abelian root numbers [4], [12]. As noted in the introduction, the proof is too long to be included in this paper.

If X is a linear character of a global Galois group, $G(K/k)$, then the global abelian root number, $W(X)$, may be written as a product of local abelian root numbers, [1, 2]:

$$W(X) = \prod_{\mathfrak{p}} R(X_{\mathfrak{p}}), \tag{15}$$

the product being over all primes of k. It should be noted that (15) need not be the only possible factorization of $W(X)$, further possibilities being covered by the following definition.

Definition 6. A function, H, defined on all multiplicative characters of local number fields is said to be a *factorization of unity* if it satisfies hypothesis (b) of Theorem 3 and if in that notation

$$1 = \prod_{\mathfrak{p}} H(X_{\mathfrak{p}}), \qquad (X \text{ linear}) \tag{16}$$

the product being over all primes of the ground field.

Theorem 3 and the pertinent parts of Theorem 1 may be reformulated so as to make Theorem 4 more directly applicable.

THEOREM 5. *Let F be a function defined on all multiplicative characters of local number fields such that*

(a) *In the notation of Theorem* 4a, $\prod_{\theta|\Theta} F(\theta) = F(\Theta)$.

(b) *In the notation of Theorem* 4b, $F(\theta_1) = F(\theta_2)$.

(c) *F satisfies conditions* (b) *and* (c) *of Theorem* 3.

---

[7] Theorem 4a in this more general form has been independently stated and partially verified by Hasse [5]. For a full proof see the dissertation referred to in footnote 1.

Then F is extendable with respect to the family of all local Galois groups. If $\bar{M}$ is the extension of M and $\bar{F}$ is the extension of F then for each character, X, of a global Galois group, $G(K/k)$,

$$\bar{M}(X) = \prod_{\mathfrak{p}} \bar{F}(X_{\mathfrak{p}}),$$

the product being over all primes of k and M being as in the notation of condition (c) *of Theorem* 3.

Proof. Let Δ be a family of nilpotent local Galois groups. We assert that F satisfies the conditions of Theorem 1 for Δ. F certainly satisfies conditions (a) and (d) of Theorem 1 and in the statement of condition (c) of that theorem $G = G_1$ and $G_0 = H$ as G is nilpotent. Hence condition (c) of Theorem 1 may for this application be written: If Θ is a linear character of H which is the restriction to H of a linear character of G then $F(\Theta) = \prod_{\theta|H=\Theta} F(\theta)$, the product being over all linear characters of G whose restriction to H is Θ. Furthermore in the statement of condition (b) of Theorem 1, as θ_i $(i = 1, 2)$ induces an irreducible character of G, it follows from the Frobenius reciprocity theorem that θ_i cannot be the restriction to G_i of a linear character of G. Also, as θ_1 and θ_2 induce the same irreducible character of G, it follows from Mackey, [8], that there exists $\sigma \varepsilon G$ such that $\theta_1{}^{\sigma}$ coincides with θ_2 on H. As $G = G_1G_2$ and H contains the commutator subgroup of G, it follows that θ_1 and θ_2 coincide on H. That hypothesis (a) of Theorem 5 implies that F satisfies condition (c) of Theorem 1 and that hypothesis (b) of Theorem 5 implies that F satisfies condition (b) of Theorem 1 now follows from a well known fact: Let B be a local number field, A a normal extension field of finite degree and C an intermediate field. If θ is a character of B^* which corresponds to a linear character, X, of $G(A/B)$ and θ_0 is the character of C^* which corresponds to the restriction of X to $G(A/C)$ then $\theta_0 = \theta \circ N_{C/B}$, the composition of θ with the relative norm of C over B.

It now follows from Theorem 1 that condition (a) of Theorem 3 is satisfied. This proves the extendability of F. The final assertion concerning the functions, $\bar{F}$ and $\bar{M}$ is a direct consequence of the uniqueness of $\bar{M}$ and the previously used argument of Artin [11], pp. 3-5.

Our main result may now be stated. It is a direct consequence of Theorems 4 and 5 and the fact that $\theta \to \theta(-1)$ is a factorization of unity.

THEOREM 5′. (a) *Let F be a function defined on linear characters of local Galois groups by setting $F(\theta) = [R(\theta)]^2\theta(-1)$ for each multiplicative character, θ, of a local number field (with the usual identification), then F*

is extendable with respect to the family of all local Galois groups. Denoting the extended function by F, $[W(X)]^2 = \prod_{\mathfrak{p}} F(X_{\mathfrak{p}})$, for each character, X, of a global Galois group, the product being over all primes of the ground field.

(b) *If H is a factorization of unity such that in the notation of*

$$\text{Theorem } 4(\text{a}),\ [\prod_{\theta|\Theta} H(\theta)]/(\Theta) = V(A/B)$$

$$\text{Theorem } 4(\text{b}),\ H(\theta_1)V(A_1/B) = H(\theta_2)V(A_2/B)$$

then $F' : \theta \to R(\theta)/H(\theta)$ defines a function on linear characters of local Galois groups which is extendable with respect to every family of local Galois groups and the extended function gives a factorization of the Artin root number in the obvious way.

4. The question of sign. The existence of a function, H, satisfying the conditions of Theorem 5(b) is still an open question and for this reason the existence of local non-abelian root numbers and the factorization of the Artin root number is established only to within factors ± 1. This aspect of the problem shall now be considered. As $V(A/B) = 1$ if degree A/B is odd (Theorem 4), it is to be expected that it is enough to determine $H(\theta)$ for θ of period 2^m (as used in this section 2^m is to be understood to be a generic integral power of 2). More precisely:

LEMMA 4. *Let J be a function defined on multiplicative characters of period 2^m of local number fields such that*

(a) *If X is a linear character of period 2^m of a global Galois group, $G(K/k)$, then $J(X_{\mathfrak{p}}) = 1$ for almost all primes of k, and $\prod_{\mathfrak{p}} J(X_{\mathfrak{p}}) = 1$, the product being over all primes of k.*

(b) *In the notation of Theorem 4(a), if $\theta|\Theta$, θ of period 2^m and $q =$ degree A/B then*

(b_1) $[J(\theta)]^q = J(\Theta)$ *if* $q \neq 2$

(b_2) $J(\theta)J(\theta\delta)/J(\Theta) = R(\delta)$ *if $q = 2$ and δ is the character of B^* which "cuts out" A over B.*

(c) *In the notation of Theorem 4(b), if θ_1 and θ_2 are of period 2^m and degree $A/B = 4$ then $J(\theta_1)V(A_1/B) = J(\theta_2)V(A_2/B)$*

and if for each multiplicative character, θ, of a local number field, $\theta = \theta'\theta''$ denotes the natural decomposition of θ into a character, θ', of odd period and a character, θ'', of period 2^m then $H : \theta \to J(\theta'')$ is a factorization of unity which satisfies the conditions of Theorem 5(b).

Proof. It is enough to show that H satisfies the condition corresponding to Theorem 4(b). In the statement of that theorem, let $p^2 =$ degree A/B, and let $\theta_i = \theta_{i1}\theta_{i2}$, $(i = 1, 2)$ be the natural decomposition of θ_i into a character, θ_{i1}, of odd period and a character, θ_{i2}, of period 2^m. It is enough to show that

$$J(\theta_{12})V(A_1/B) = J(\theta_{22})V(A_2/B). \tag{17}$$

Let σ be a non-trivial element of the Galois group, $G(A/A_2) \cong G(A_1/B)$, then since $\theta_1 \circ N_{A/A_1} = \theta_2 \circ N_{A/A_2}$, it follows that $\theta_1{}^{\sigma-1} \circ N_{A/A_1}$ is the principal character of A^*. Hence $\theta_1{}^{\sigma-1}$ is trivial on $N_{A/A_1}A^*$ but is not trivial on $A_1{}^*$ as otherwise (since A_1 is cyclic over B) θ_1 is divisible by a character of B^*, contrary to hypothesis. Hence $\theta_1{}^{\sigma-1}$ is of period p. Thus $\theta_{12}{}^{\sigma-1}$ is trivial if and only if $p \neq 2$. It is clear that θ_{11} and θ_{21} divide the same character of A^* and that the same holds for characters, θ_{12} and θ_{22}. If p is odd then θ_{12} is divisible by a character, θ, of period 2^m, of B^* and by an easy argument $\theta | \theta_{22}$, whence (17) follows from (b_1). If $p = 2$ then θ_{12} is not divisible by any character of B^* so that (17) follows from (c).

A partial solution for J may be given.

LEMMA 5. *If θ is a real multiplicative character of a local number field let $J(\theta) = R(\theta)$. J satisfies the conditions of the last lemma for characters of this type.*

Proof. If X is a linear character of period 2^m of a global Galois group, $G(K/k)$, such that $X_\mathfrak{p}$ is real for each prime $\mathfrak{p}$, of k, then X is of period 2 or 1. In either case $1 = W(X) = \prod_\mathfrak{p} R(X_\mathfrak{p})$, so that condition (a) of the last lemma is satisfied. Condition (b_2) and (c) follow from Theorem 4, and for condition (b_1) let C be the quadratic extension of B cut out by θ, then AC is the extension of A cut out by Θ and AC is an abelian extension of B. Hence $R(\Theta) = V(AC/A)$, $R(\theta) = V(C/B)$ and by (14),

$$V(AC/A)[V(A/B)]^2 = V(AC/B) = V(AC/C)[V(C/B)]^q,$$

whence (c_1) follows from $V(A/B) = V(AC/C) = 1$.

It follows from this lemma that if Y is a character of a local Galois group, G, then the (local) root number of Y is well defined if no cyclic group of order 4 is in the family generated by G. Furthermore if X is a character of a global Galois group, G, then the Artin root number, $W(X)$, has a well defined decomposition if the decomposition subgroups of G satisfy this condition.

Theorem 5′ indicates that the determination of H requires a canonical

extraction of the square root of $\theta(-1)$. If θ is of period 2 then this is provided by $R(\theta)$.

We shall now show that in a certain sense it is impossible to find a function H which satisfies the conditions of Theorem 5′b. If H were found then for each multiplicative character, θ, of a local number field, B, $R(\theta)/H(\theta)$ could be viewed as a new local abelian root number. It would therefore be expected that H would satisfy the further conditions:

(I) $H(\theta)=1$ if θ is the principal character of B^*

(II) $H(\theta)=1$ if θ is an absolutely unramified character (i. e. if the kernel of θ determines a cyclic extension of B which is absolutely unramified).

It will be shown that these conditions are incompatible with the previous conditions on H. Suppose that there exists an H satisfying all the conditions. It follows from (I) that if θ is real then $H(\theta)=R(\theta)$. In particular, if B is the field of real numbers then $1/H(\theta)=\sqrt{[\theta(-1)]}$, where $\sqrt{1}=1$, $\sqrt{(-1)}=i$. If n is any positive integer, let G_n be the Galois group of $Q(\zeta_n)$ over Q, where ζ_n is a primitive n-th root of unity. If p is an odd rational prime number and X is a character of G_p then (from II), $H(X_\infty)H(X_p)=1$. Likewise if q is an odd prime and Y is a character of G_q then $H(Y_\infty)H(Y_q)=1$, while if XY is the character of G_{pq}, $(p\neq q)$, obtained in the obvious way then $H(X_\infty Y_\infty)H(X_pY_p)H(X_qY_q)=1$. As X_q is unramified and $X_q(q)X_p(q)=1$, X_q is trivial if $X_p(q)=1$, which is certainly true if $q\equiv 1 \pmod{p}$. If q satisfies this condition then

$$H(X_pY_p)=\sqrt{[X_p(-1)Y_q(-1)]}/\sqrt{[Y_q(-1)]}.$$

Given p, X may be chosen so that $X_p(-1)=-1$, in which case $H(X_pY_p)=\sqrt{[-Y_q(-1)]}/\sqrt{[Y_q(-1)]}$. Hence to obtain a contradiction, it is enough to choose p and an unramified character θ of $Q_p{}^*$ such that $Y_q(-1)$ is not independent of the choice of q and Y which satisfy the conditions: $q\equiv 1 \pmod{p}$; $Y_p=\theta$ (i.e. $1/\theta(p)=Y_q(p)$). To do this let $p=5$, $\theta(5)=i$. For $q=41$, p is a quadratic but not a biquadratic residue mod q, likewise for $q=241$. Hence in either case there exists Y of period 8 such that $Y_q(p)=-i$. As -1 is an eighth power mod 241 but not mod 41, $Y_q(-1)=-1$ for $q=41$, $Y_q(-1)=1$ for $q=241$. The contradiction is thus proven.

This counter example almost assures us of the nonexistence of a function H satisfying the conditions of Theorem 5′b. The problem of finding H is that of reformulating the definition of local abelian root numbers so that

the group theoretical properties of the Artin root number may be explained locally. Instead we can consider the global relations between local abelian root numbers which can be deduced from the properties of the Artin root number. Let K be a normal overfield of an algebraic number field, k; let ϕ_i $(i=1,\cdots,r)$ be a linear character of a subgroup, $G(K/k_i)$ of $G(K/k)$ such that $\sum_{i=1}^{r} a_i X_i = 0$, where X_i is the character of $G(K/k)$ which is induced by ϕ_i and a_i is an integer. The global relations referred to above are of the form

$$\prod_{\mathfrak{p}} \prod_{i=1}^{r} \prod_{\mathfrak{q}_i \mathfrak{p}} R((\phi_i)_{\mathfrak{q}_i})^{a_i} = 1$$

($\mathfrak{p}$ runs through all primes of k; $\mathfrak{q}_i$ runs through all primes of k_i). Following Hasse, let δ_i be the character of $G(K/k)$ induced by the principal character of $G(K/k_i)$. Let $|\delta_i|$ denote the linear character of $G(K/k)$ obtained by taking the determinant of the matrix representation corresponding to δ_i. As a direct consequence of Theorem 5′a, the $\mathfrak{p}$ constituent of the above relation is given by

$$\prod_{i=1}^{r} \prod_{\mathfrak{q}_i | \mathfrak{p}} R((\phi_i)_{\mathfrak{q}_i})^{a_i} = \pm \prod_{i=1}^{r} [R(|\delta_i|_{\mathfrak{p}})]^{-a_i}, \qquad \text{(cf. [5, page 83])}$$

where the inner product on the left is over all primes of k_i which divide $\mathfrak{p}$. As $|\delta_i|$ is either the principal character or a quadratic character of $G(K/k)$, it is natural to conjecture that all of these global relations are products of two kinds of relations:

1. If X is any linear character of $G(K/k)$ then $\prod_{\mathfrak{p}} X_{\mathfrak{p}}(-1) = 1$.

2. If X is a quadratic character of $G(K/k)$ then $\prod_{\mathfrak{p}} R(X_{\mathfrak{p}}) = 1$.

(The product in both cases is over all primes of k.)

5. Local root numbers. In this section let C^* be the multiplicative group of non-zero complex numbers and let Γ be the factor group, $C^*/\{\pm 1\}$. If X is a character of a local Galois group, $G = G(K/k)$, let $W(X)$ denote the *local root number* $(\varepsilon \Gamma)$ which is obtained by extending the function, $X \to R(X) \vee [X(-1)]$, defined on linear characters of local Galois groups (Theorem 5′(a)). It follows from Definition 5 that if k is Archimedean then $W(X) = 1$, the identity element of Γ. Hence in studying local root numbers it may be assumed that k is a $\mathfrak{p}$-adic number field. In the $\mathfrak{p}$-adic case we define (following Hasse) the *local Galois Gauss* sum, $\tau(X) = W(X) \vee [Nf(X)]$,

where $Nf(X)$ is the absolute norm of the conductor of X. If Y is a character of a global Galois group, $G(E/F)$, then Hasse's Galois Gauss sum, $\tau(Y)$, is defined by

$$\tau(Y) = W(Y) \sqrt{[Nf(Y)]}\, \varepsilon \prod_{p} \tau(Y_p), \tag{18}$$

the product over all primes of F being understood to be a coset of $\{\pm 1\}$. In the following the symbols, $W(X)$, $\tau(X)$, shall refer to elements of C^* if X is a character of a global Galois group and elements of Γ (or cosets of $\{\pm 1\}$) if X is a character of a local Galois group. An element of Γ will be said to be a root of unity (resp.: an algebraic integer, resp.: an element of a certain algebraic number field) if the elements of the corresponding coset of $\{\pm 1\}$ have this property.

Continuing with the discussion of the local situation, let $\mathfrak{W}_0$ be the inertial subgroup, $\mathfrak{W}_1, \mathfrak{W}_2, \cdots$ the ramification subgroups of G. Let

$$m(X) = [\mathfrak{W}_0]^{-1} \sum_{i=0}^{s} [\mathfrak{W}_i], \tag{19}$$

where s is the smallest integer such that X is trivial (i.e. identically $X(1)$) on $\mathfrak{W}_{s+1}$, $m(X)$ being understood to be zero if $s = -1$. Letting $\mathfrak{p}$ be the prime ideal of k, the conductor, $f(X)$, is given by Artin [11],

$$\operatorname{ord}_{\mathfrak{p}} f(X) = [\mathfrak{W}_0]^{-1} \sum_{i=0}^{\infty} \{[\mathfrak{W}_i] X(1) - X(\mathfrak{W}_i)\}, \tag{20}$$

where $X(\mathfrak{W}_i) = \sum_{w \in \mathfrak{W}_i} X(w)$, this last sum being over all elements of $\mathfrak{W}_i$.

LEMMA 6. *If X is an irreducible character of G then*

$$m(X) = [\operatorname{ord}_{\mathfrak{p}} f(X)]/X(1) \tag{21}$$

and if $X = \sum_{j=1}^{r} X_j$, a sum of characters of G then

$$m(X) = \operatorname{Max}_{1 \leq j \leq r} m(X_j). \tag{22}$$

Proof. From the orthogonality relations, $n_i = [\mathfrak{W}_i]^{-1} X(\mathfrak{W}_i)$ is the number of times the trivial character of $\mathfrak{W}_i$ occurs in $X | \mathfrak{W}_i$. If X is irreducible then $X | \mathfrak{W}_i$ is a sum of conjugate characters of $\mathfrak{W}_i$ as $\mathfrak{W}_i$ is an invariant subgroup of G. Hence n_i is either $X(1)$ or zero depending upon whether or not i exceeds s. Thus (21) follows from (19) and (20). If $X = \sum X_j$ and X is trivial on $\mathfrak{W}_i$ then each X_j is trivial on that group, i.e. $m(X) \geqq m(X_j)$. Conversely if each X_j $(j = 1, \cdots, r)$ is trivial on $\mathfrak{W}_i$ then X is trivial on $\mathfrak{W}_i$, i.e. $\operatorname{Max}_j m(X_j) \geqq m(X)$ which proves (22).

It follows from the definition that either $m(X) = 0$ or $m(X) \geqq 1$. It

is easily shown by the methods of Artin, [11, page 9], that $m(X)$ is a rational number whose denominator is a power of the rational prime, p, which is divisible by $\mathfrak{p}$. $m(X)$ need not be an integer, a counter example being provided by the situation discussed in Theorem 4(b). In the notation of that theorem let C be the cyclic extension of A which is cut out by the character, $\theta_1 \circ N_{A/A_1}$, of A^*. Let A be totally and wildly ramified over B and let θ_1 be chosen so as to have the "minimal" conductor (i.e. with smallest possible exponent) and let X be the (irreducible) character of $G(C/B)$ which is induced by the linear character of $G(C/A_1)$ which corresponds to θ_1. Under these circumstances $m(X)$ is not an integer but the details of the computation are not needed at this time.

We now show that if X is irreducible then $m(X)$ determines the nature of the local root number, $W(X)$.

THEOREM 6. *Let X be an irreducible character of the local Galois group, $G(K/k)$.*

(a) *If $m(X) \leqq 1$ then there exists an intermediate field, k', unramified over k such that X is induced by a linear character, Z, of $G(K/k')$ of conductor $\mathfrak{q}^{m(X)}$ ($\mathfrak{q}$ = prime of k') and $\tau(X) = \tau(Z)$, hence an algebraic integer.*

(b) *If $m(X) \neq 1$ then $W(X)$ is a root of unity.*

Proof. If $m(X) = 0$ then X is trivial on $\mathfrak{W}_0$, whence X is linear and the theorem is trivial in that case. If $m(X) = 1$ then X is trivial on $\mathfrak{W}_1$, whence it is induced by a linear character, Z, of a subgroup, G', of G which contains $\mathfrak{W}_0$ (as $\mathfrak{W}_0/\mathfrak{W}_1$ is cyclic, hence certainly abelian). G' corresponds to a subfield, k', of K which is unramified over k, so that $N_{k'/k} f(Z) = f(X)$, whence $m(Z) = 1$. Clearly $Nf(X) = Nf(Z)$ so that $\tau(X) = \tau(Z)$ which completes the proof of (a).

If $m(X) > 1$ then there exists a minimal integer, s, $s > 0$, such that X is trivial on $\mathfrak{W}_{s+1}$. As $\mathfrak{W}_s/\mathfrak{W}_{s+1}$ is abelian, X is induced by a character, X', of a subgroup, G', of $G(K/k)$ which contains $\mathfrak{W}_s$ and leaves δ fixed, δ being one of the linear characters of $\mathfrak{W}_s$ which occurs in $X|\mathfrak{W}_s$ (Lemma 2). For the same reason, X' has a Brauer decomposition in characters of G' which are induced by linear characters, Y_j, of subgroups, G_j, of G' which contain $\mathfrak{W}_s$, such that $\delta = Y_j|\mathfrak{W}_s$. δ is a non-trivial character of $\mathfrak{W}_s$ as $X|\mathfrak{W}_s$ contains only conjugates of δ, hence the conductor of δ is $\mathfrak{P}^r$, $r > 1$, $\mathfrak{P}$ being the prime of V_s, the s-th ramification subfield of K over k. If k_j is the subfield of K which corresponds to G_j then $V_s \supset k_j \supset k$ and Y_j corresponds to a linear character, θ_j, of k_j^*. Also $\theta_j \circ N_{V_s/k_j} = \Delta$, the character of V_s^* which corresponds to δ. As $N_{V_s/k_j}(1 + \mathfrak{P}) \subset 1 + \mathfrak{p}_j$ ($\mathfrak{p}_j$ = prime of k_j), if $\mathfrak{p}_j^2 \nmid f(\theta_j)$

then $\mathfrak{P}^2 \nmid f(\Delta) = f(\delta)$, a contradiction. Hence $\mathfrak{p}_j^2 | f(\theta_j)$ so that the abelian root number, $R(\theta_j)$, is a root of unity. As $W(X)$ is a product of such quantities and their inverses, and powers of i, (b) follows from the previous discussion of the case $m(X) = 0$.

As an immediate consequence of this theorem and equation (18) we have a conjecture of Hasse:

COROLLARY. *If X is a character of a global Galois group then the Galois Gauss sum, $\tau(X)$, is an algebraic integer.*

We shall now consider Hasse's conjecture concerning the algebraic number field in which the Artin root number lies. It follows from the definitions that if X is a linear character of a local Galois group, $G(K/k)$, k being a $\mathfrak{p}$-adic number field, then $[W(X)]^2 = [R(X)]^2 X(-1)$ lies in the field, A, obtained by adjoining to Q (= the field of rational numbers) the values assumed by X and a primitive N-th root of unity, where N is the power of p (= the rational prime divisible by $\mathfrak{p}$) defined by the ideal theoretic relation

$$(23) \qquad 1/N = S_{k/Q_p}(1/(\mathfrak{D}_k f(X))),$$

Q_p being the p-adic completion of Q. $W(X)$ itself lies in the field obtained by adjoining to A the elements i and $\sqrt{p}$. The generalization of this result is complicated by the uncertain status of the question of sign (Section 4).

THEOREM 7. *If X is a character of a local Galois group, $G(K/k)$, k being a $\mathfrak{p}$-adic number field, then $[W(X)]^2$ lies in the field $A = Q(X; m(X)/e)$ obtained by adjoining to Q the values assumed by X and the N roots of unity, where $N = p^r$, p is the rational prime divisible by $\mathfrak{p}$, r is the smallest integer not less than $m(X)/e$, e being the absolute ramification of k.*

Proof. The theorem is first proven for X irreducible. If X is linear then the assertion follows from the fact that $S_{k/Q_p}(1/(\mathfrak{D}_k f(X))) = p^{[-m(X)/e]}$, the brackets denoting the largest integer not greater than $-m(X)/e$. If $m(X) = 1$ then by Theorem 6, $W(X) = W(Z)$, Z being a linear character of $G(K/k')$ where k' is an unramified extension of k. As $m(Z) = 1$ the assertion follows from the discussion of the linear case. If $m(X) > 1$ then in the notation of the proof of the preceding theorem, $[W(X)]^2$ lies in the composition of fields of the type $Q(Y_j; m(Y_j)/e_j)$, Y_j being a linear character of $G_j = G(K/k_j)$ which contains $\mathfrak{W}_s$, e_j being the absolute ramification of k_j. Furthermore $Y_j | \mathfrak{W}_s = \delta$. As the i-th ramification subgroup of G_j is $\mathfrak{W}_i \cap G_j$,

$\mathfrak{W}_s$ is the s-th ramification subgroup of G_j and $\mathfrak{W}_{s+1}$ is the $(s+1)$-th ramification subgroup of G_j. Y_j is trivial on $\mathfrak{W}_{s+1}$ but not on $\mathfrak{W}_s$. Hence

$$m(Y_j)/e_j = (e_j[\mathfrak{W}_0 \cap G_j])^{-1} \sum_{i=0}^{s} [G_j \cap \mathfrak{W}_i]. \tag{24}$$

But $e_j[\mathfrak{W}_0 \cap G_j] =$ absolute ramification of $K = e[\mathfrak{W}_0]$ so that $m(Y_j) e_j \leqq m(X)/e$. It follows that if $Y_1, \cdots, Y_t$ are the linear characters of subgroups of G which are involved in the Brauer decomposition of X described in the proof of Theorem 6 then $[W(X)]^2 \varepsilon\, Q(Y_1, \cdots, Y_t; m(X)/e) = B$, it being understood that the field, B, contains all values assumed by $Y_1, \cdots, Y_t$ respectively. Certainly $B \supset A$, but A contains all roots of unity introduced by the additive characters of the fields $k_1, \cdots, k_t$ in forming the abelian Gauss sums associated with the linear characters $Y_1, \cdots, Y_t$. The Brauer decomposition of X may be written

$$X = \sum_{j=1}^{t} a_j X_j, \tag{25}$$

X_j being the character of G which is induced by Y_j, and each a_j being an integer. Clearly, $W(X) = \prod_{j=1}^{t} W(Y_j)^{a_j}$. If σ is an isomorphism of B which leaves the elements of A fixed then it leaves each value of X fixed so that (25) may be written with each X_j replaced by the character $\sigma \circ X_j$ which is induced by $Y_j^\sigma = \sigma \circ Y_j$. Hence $W(X) = \prod_{j=1}^{t} W(Y_j^\sigma)^{a_j}$. But σ leaves fixed the values assumed by the additive characters, i. e. $[W(Y_j^\sigma)]^2 = [W(Y_j)^\sigma]^2$. It follows that σ leaves $[W(X)]^2$ invariant, whence $[W(X)]^2$ lies in A.

If $X = \sum_i L_i$, a sum of irreducible characters of G then by the preceding argument and (22) it follows that $[W(X)]^2 \subset Q(L_1, \cdots, L_r, m(X)/e) = A'$. The same argument as in the irreducible case shows that if σ is an isomorphism of A' which leaves the elements of A fixed then

$$[W(X)]^2 = [\prod_i W(L_i^\sigma)]^2 = [W(X)^\sigma]^2$$

as σ leaves invariant all values assumed by additive characters used in expressing $W(L_i)$ in terms of local abelian root numbers.

This result has an obvious global consequence.

COROLLARY 1. *If X is a character of a global Galois group, $G(K/k)$, then $[W(X)]^2$ lies in the field A obtained by adjoining to Q the values assumed by X and a primitive N-th root of unity, where*

$$\log N = \sum_p (\log p) \max_{\mathfrak{p} | p} m_{\mathfrak{p}}, \tag{27}$$

$m_{\mathfrak{p}}$ being the smallest integer which is not less than $m(X_{\mathfrak{p}})/e_{\mathfrak{p}}$, $e_{\mathfrak{p}}$ being the absolute ramification of $\mathfrak{p}$, the sum being over all finite rational primes.

The following corollaries refer to both local and global characters. The notation is as in Theorem 7 in the local case and as in Corollary 1 in the global case. If n is an integer let Z_n denote a primitive n-th root of unity.

COROLLARY 2. *$W(X)$ lies in $Q(Z_r)$, where r is the least common multiple of 8, N and the degree, n, of K over k.*

Proof. In the local case it follows by the same reasoning as in Theorem 7 (without the automorphism argument as the values assumed by the characters, Y_j, are n-th roots of unity) that $W(X) \varepsilon Q(Z_n, Z_N, i, \sqrt{p})$. If $m(X) = 0$ then $W(X)$ is an n-th root of unity, while if $m(X) > 0$ then $Z_p \varepsilon Q(Z_N)$. As i, $\sqrt{p}$, lie in $Q(Z_p, i)$ for $p \neq 2$, and i, $\sqrt{2}$ lie in $Q(Z_8)$ the assertion follows in the local case. The proof in the global case is now immediate.

COROLLARY 3. *If the question of sign is solved for the local Galois groups involved with $G(K/k)$ (in particular if the family generated by $G(K/k)$ contains no cyclic group of order 4) then $W(X) \varepsilon A(Z_8)$.*

The proof is a repetition of previous arguments. We note that everything said here for root numbers (i.e. in Theorem 7 and its corollaries) applies equally well to Galois Gauss sums.

Final Remarks. The proof of the integrality of Galois Gauss sums requires only a weak form of Theorem 5′. It is enough to know the extendability of the local abelian root number considered as a function which takes its values in the group C^*/E, where E is the group of all roots of unity in the complex field. It follows that a weak form of Theorem 4 is adequate, it being enough to know that (10) and (12) are valid as congruences modulo E. These congruences are a trivial consequence of the multiplicative identities of [12] and in fact as a congruence, (12) is a direct consequence of (10).

The results of this paper may be extended without difficulty to the root number in the functional equation of the L-series studied by Weil in [14]. This may be done most easily by using the fact that an irreducible character of one of Weil's decomposition groups is the product of a linear unramified character and an irreducible character with kernel of finite index and the latter may be identified with a character of a local Galois group. While the root numbers appearing in this theory need not be algebraic numbers, Theorem 6 may be rewritten so as to remain valid.

Finally it may be noted that the methods of this paper may be used to prove the existence and most of the properties of the Artin conductor, [11], without the use of Artin's specific formula.

HARVARD UNIVERSITY.

REFERENCES.

[1] E. Artin, "Zur Theorie der L-Reihen mit allgemeinen Gruppencharakteren," *Abhandlungen aus dem Mathematischen Seminar der Universität Hamburg*, vol. 8 (1931), pp. 292-306.

[2] J. Tate, "Fourier Analysis in Number Fields and Hecke's Zeta Functions," Thesis, Princeton University, 1950 (unpublished).

[3] H. Hasse, "Allgemeine Theorie der Gaussche Summen in algebraischen Zahlkorpern," *Abhandlungen der Deutschen Akademie der Wissenschaften zu Berlin*, Math-natur-wiss. Kl. Jahrg. 1951, Nr. 1, pp. 1-23.

[4] E. Lamprecht, "Allgemeine Theorie der Gaussche Summen in endlichen kommutativen Ringen," *Mathematische Nachrichten*, vol. 9 (1935), pp. 149-196.

[5] H. Hasse, "Artinsche Führer, Artinsche L-Funktionen und Gaussche Summen uber Endlich-Algebraischen Zahlkorpern," *Acta Salmanticensia*, Ciencias: Seccion de Mathematicas, 1954.

[6] R. Brauer, "On Artin's L-series with general group characters," *Annals of Mathematics*, vol. 48 (1947), pp. 502-514.

[7] E. Artin, "Galois Theory" (second edition), Notre Dame Mathematical Lectures, Notre Dame, Indiana (1946).

[8] G. Mackey, "On induced representations of groups," *American Journal of Mathematics*, vol. 73 (1951), pp. 576-592.

[9] K. Taketa, "Uber die Gruppen deren Darstellung samtlich auf monomiale gestalt transformieren lassen," *Proceedings of the Imperial Academy of Japan*, vol. 6 (1930).

[10] E. Artin, "Algebraic numbers and algebraic functions," vol. I, Princeton University, 1950-1951.

[11] ———, "Die gruppentheoretische Struktur der Diskriminanten algebraischer Zahlkörper," *Journal für die Reine und Angewandte Mathematik*, vol. 164 (1931), pp. 1-11.

[12] H. Davenport and H. Hasse, "Die Nullstellen der Kongruenzzetafunktionen in gewissen zyklischen Fällen," *Journal für die Reine und Angewandte Mathematik*, vol. 172 (1934), pp. 151-182.

[13] B. Huppert, "Normalteiler und maximale Untergruppen endlicher Gruppen," *Mathematische Zeitschrift*, vol. 60 (1954), pp. 409-434.

[14] A. Weil, "Sur la théorie du corps de classes," *Journal of the Mathemical Society of Japan*, vol. 3 (1951), pp. 1-35.

ARTIN-SCHREIER EQUATIONS IN CHARACTERISTIC ZERO.*

By R. E. MacKenzie and G. Whaples.

The proof of the existence theorem of generalized local class field theory [6, 7] shows that if k is a field over which that theory holds and p is the characteristic of its residue class field then the cyclic extensions of k of degree p fall into two disjoint classes. Each extension of the first class is contained in the composite of finitely many extensions generated by root of Artin-Schreier equations

$$(1) \qquad x^p - x - \lambda = 0, \qquad \lambda \varepsilon k, \qquad | p^p \lambda^{p-1} | < 1.$$

Each extension of the second class is generated by a root of an equation

$$(2) \qquad x^p - \alpha = 0, \qquad \alpha \varepsilon k, \qquad | \alpha | \not\varepsilon | k^{\cdot p} |,$$

where $| k^{\cdot p} | = \{ | \beta^p | \mid \beta \varepsilon k^{\cdot} \}$. Extensions of the second class occur only when k has characteristic zero and contains primitive p-th roots of unity.

This suggests the conjecture that every extension of the first class is generated by a root of *one* Artin-Schreier equation. This is well known, of course, when k has characteristic p [9]. In this paper we prove the conjecture not only for the fields of local class field theory but for all fields which are maximally complete [2, 4] under an arbitrary non-archimedean valuation (not necessarily discrete or of rank one) and which have a residue class field with no inseparable extension. Even without this assumption on the residue class field our methods give considerable information about the cyclic extensions of degree p.

In local class field theory the two classes of extensions can be distinguished by certain inequalities involving either the conductor or the different. In our more general situation the conductor and different are unavailable but we introduce another invariant, the distortion constant, which takes their place and happens also to simplify the computations very much.

This gives a method of assigning to each cyclic extension of degree p a canonical defining equation. In the local class field case this poses the

* Received May 26, 1954; revised February 3, 1956. Work of Whaples on this paper supported in part by National Science Foundation Grant NSF G1916.

problem of finding an explicit reciprocity law, namely, a rule translating the class field theory parametrization of extension fields and their automorphisms into the parametriaztion given by our canonical defining equations. Such laws have up to now been restricted to the case where k has characteristic p or contains primitive p-th roots of unity due to the lack of a canonical defining equation in the other case. The problem of finding an explicit reciprocity law separates naturally into two parts: (A) Find an explicit formula for the norm residue symbol defined on some particular basis for elements of $k^{\cdot}$ modulo norms. (B) Find an explicit basis for the group of norms. We have solved (A) but not (B).

Oystein Ore [3] has also studied generating equations for such cyclic extensions (not only of degree p but of degree p^n) in the absence of p-th roots of unity. Although his general ideas are similar to ours he restricts himself to algebraic number fields, uses congruences instead of equations, and, because he wants congruences with integral coefficients, has several standard forms instead of only two.

For brevity we state at the beginning of each section the assumptions made in it and do not repeat these assumptions in stating propositions and theorems.

1. Orthobases. Let K have a nonarchimedean valuation $|\ |$ with no assumptions of completeness. Let K/k be finite algebraic and $\bar{K}/\bar{k}$ the residue class field extension. Let n, e, f denote the degree of K/k, the ramification number of K/k, and the degree of $\bar{K}/\bar{k}$, respectively.

Definition 1. A sequence $A_1, A_2, \cdots, A_n$ of elements of K is called an *orthobasis for* K/k (relative to $|\ |$) if it is a basis and

$$|\textstyle\sum_\nu \alpha_\nu A_\nu| = \max_\nu |\alpha_\nu A_\nu| \text{ for all } \alpha_1, \alpha_2, \cdots, \alpha_n \in k. \tag{3}$$

The name was chosen because, like an orthonormal basis of a normal vector space, an orthobasis makes absolute values easy to compute. If the valuation is not discrete there may not exist any minimal basis for K-integers over k-integers, but an orthobasis is an excellent substitute.

PROPOSITION 1. *K/k has an orthobasis if and only if $ef = n$.*

Proof. Let $\{A_i\}$ be a finite set of elements of K such that $\{|A_i|\}$ are in distinct cosets modulo $|k^{\cdot}|$ and let $\{B_j\}$ be elements of value 1 whose residue classes are linearly independent over $\bar{k}$. It is easy to see that the elements $\{A_iB_j\}$ satisfy (3) and are therefore linearly independent over k. The well known statement $n \geqq ef$ [1] follows. If $n > ef$ then any set of n

elements of K contains more than f elements whose values are in the same coset modulo $|k^{\cdot}|$ and hence fails to satisfy (3).

Definition 2. An element $A \varepsilon K$ is called an *orthogenerator for K/k* if $1, A, A^2, \cdots, A^{n-1}$ are orthobasis.

2. Distortion constant. Besides the assumptions of Section 1 let K/k be cyclic with a generating automorphism σ such that

$$|\sigma A| = |A| \text{ for all } A \varepsilon K. \tag{4}$$

Definition 3. An element $\Gamma \varepsilon K$ is called a *distortion constant* (d.c.) for K/k when

$$\sigma A = A(1+\Gamma) \text{ for some } A \varepsilon K \text{ and} \tag{5}$$

$$\sigma B = B(1+O(\Gamma)) \text{ for all } B \varepsilon K, \tag{6}$$

where $O(\Gamma)$ denotes an element of value $\leqq |\Gamma|$.

Clearly K/k has a d.c. if and only if $((\sigma-1)B/B = B^{\sigma-1}-1$ assumes a maximum value for some $B = A \varepsilon K$. If any d.c. exists the value of every d.c. equals this maximum, hence is an invariant of K/k and σ. Since $\Gamma = ((\sigma-1)A)/A = A^{\sigma-1}-1$ it follows from (4) that $|\Gamma| \leqq 1$. By induction using (5) we obtain for all i

$$\sigma^i A = A(1+\Gamma)(1+\sigma\Gamma)\cdots(1+\sigma^{i-1}\Gamma) = A(1+O(\Gamma)). \tag{7}$$

Similarly from (6) we see that $\sigma^i B = B(1+O(\Gamma))$ for all i and all B. If τ is another generating automorphism and Γ' a d.c. corresponding to it then $\Gamma' = O(\Gamma)$ and $\Gamma = O(\Gamma')$ so $|\Gamma|$ does not depend on the choice of σ and is an invariant of K/k.

The essential idea of a d.c. and of the following proposition appears in O. F. G. Schilling's book [4, pp. 80, 81].

PROPOSITION 2. *If K/k has an orthobasis $A_1, A_2, \cdots, A_n$ and $A_i^{\sigma-1}-1$ is of maximal value for $i = 1, 2, \cdots, n$ then it is a d.c. If A is an orthogenerator then $A^{\sigma-1}-1$ is a d.c.*

Proof. Let $A_1, A_2, \cdots, A_n$ be an orthobasis and let $\Gamma = A_i^{\sigma-1}-1$ be of maximal value. Then for all $B = \sum \alpha_\nu A_\nu \varepsilon K$ we have $\sigma B = \sum \alpha_\nu A_\nu (1+O(\Gamma))$ so $|(\sigma-1)B| \leqq \max_\nu |\alpha_\nu A_\nu O(\Gamma)| \leqq |B\Gamma|$ by (3). Hence $\sigma B = B(1+O(\Gamma))$ and Γ is a d.c. If A is an orthogenerator and $\sigma A = A(1+\Gamma)$ then $\sigma A^\nu = A^\nu(1+\Gamma)^\nu = A^\nu(1+O(\Gamma))$ because $|\Gamma| \leqq 1$. So from what was just proved Γ is a d.c.

3. Cyclic extensions of degree p. Besides the assumptions of Sections 1 and 2 assume K/k is cyclic of degree p where $0 \leqq |p| < 1$ and $ef = p$. p is the characteristic of $\bar{k}$ and may or may not be that of k. Our results are of interest only in the case when k has characteristic 0.

Then K/k always has an orthogenerator A.

Definition 4. If $e = p$ let A be any element with $|\mathrm{A}| \not\in |k^{\cdot}|$. If $f = p$ and $\bar{K}/\bar{k}$ is inseparable let A be any element of value 1 whose residue class is not in $\bar{k}$. If $f = p$ and $\bar{K}/\bar{k}$ is separable choose A of value 1 so that $\sigma\bar{\mathrm{A}} = \bar{\mathrm{A}} + 1$, i.e., $\sigma \mathrm{A} = \mathrm{A} + 1 + o(1)$. This is possible because $\bar{k}$ has characteristic p [8]. Call an A chosen in this way a *canonical orthogenerator* of K/k.

K/k has a d.c. Γ, and $|\Gamma| < 1$ unless $\bar{K}/\bar{k}$ is separable of degree p. We propose now to obtain sharp estimates of the relative values of B, $(\sigma - 1)\mathrm{B}$, and $S\mathrm{B}$ of the sort usually obtained by using the different.

Let A be a canonical orthogenerator and $\sigma\mathrm{A} = \mathrm{A}(1+\Gamma)$. Then $(\sigma-1)\mathrm{A}^i = \mathrm{A}^i((1+\Gamma)^i - 1)$. If $|\Gamma| < 1$ then $(1+\Gamma)^i - 1 = i\Gamma + o(\Gamma)$ and has value equal to $|\Gamma|$ when $p \nmid i$. If $|\Gamma| = 1$ then, according to definition 4, $\bar{\Gamma} = \bar{\mathrm{A}}^{-1}$ so $\bar{\Gamma}$ generates $\bar{K}/\bar{k}$ and $|(1+\Gamma)^i - 1| = 1 = |\Gamma|$ when $0 < i < p$. This proves

$$(8) \qquad |(\sigma-1)\mathrm{A}^i| = |\Gamma \mathrm{A}^i| \text{ for } 1 \leqq i \leqq p-1$$

in all cases. Consider the elements

$$(9) \qquad 1, \Gamma^{-1}(\sigma-1)\mathrm{A}, \Gamma^{-1}(\sigma-1)\mathrm{A}^2, \cdots, \Gamma^{-1}(\sigma-1)\mathrm{A}^{p-1}.$$

If $|\Gamma| < 1$ then the previous remarks show that these element are equal to $1, \mathrm{A} + o(\mathrm{A}), 2\mathrm{A}^2 + o(A^2), \cdots, (p-1)\mathrm{A}^{p-1} + o(\mathrm{A}^{p-1})$, respectively, and hence form an orthobasis. If $|\Gamma| = 1$ then $\bar{K}/\bar{k}$ is separable of degree p and $\bar{\Gamma} = \bar{\mathrm{A}}^{-1}$. The residue classes of the elements (9) are $\bar{\mathrm{A}}^{i+1}((1+\bar{\mathrm{A}}^{-1})^i - 1)$ for $i = 1, \cdots, (p-1)$. They form a basis of $\bar{K}/\bar{k}$ so the elements (9) form an orthobasis of K/k. It therefore follows that the elements

$$(10) \qquad \Gamma, (\sigma-1)\mathrm{A}, (\sigma-1)\mathrm{A}^2, \cdots, (\sigma-1)\mathrm{A}^{p-1}$$

form an orthobasis in all cases.

Let B be any element of K. If we express B in terms of the orthobasis (10) we obtain $\mathrm{B} = \beta\Gamma + \Delta$ where $\beta \in k$, $S\Delta = 0$, and $|\beta\Gamma| \leqq |\mathrm{B}|$. Since $S\mathrm{B} = \beta S\Gamma$ we obtain

$$(11) \qquad |S\mathrm{B}| \leqq |\Gamma^{-1}S\Gamma| \, |\mathrm{B}|$$

and equality holds only if $|\beta\Gamma| = |\mathrm{B}|$, that is, $|\mathrm{B}| \in |\Gamma| \, |k^{\cdot}|$.

We may also express B in terms of the powers of a canonical orthogenerator A; $B = \alpha_0 + \alpha_1 A + \cdots + \alpha_{p-1} A^{p-1}$. Then $(\sigma - 1)B$ is expressed in terms of the orthobasis (10). Using (8) we obtain

$$(12) \qquad |(\sigma - 1)B| \leqq |\Gamma B|$$

and if we put $B' = B - \alpha_0$ we can say there is a B' such that

$$(13) \qquad (\sigma - 1)B' = (\sigma - 1)B \text{ and } |B'| = |\Gamma^{-1}(\sigma - 1)B|.$$

THEOREM 1. $|\Gamma^{p-1}| \geqq |p|$ *for all* K/k *satisfying the assumptions of Section* 3.

Proof. Let A be a canonical orthogenerator and Γ its d.c. We can assume $|\Gamma| < 1$. By (7) we see that $\sigma^p A = AN(1 + \Gamma)$ so $N(1 + \Gamma) = 1$, i.e. $S\Gamma + E_2 + \cdots + E_{p-1} + N\Gamma = 0$ where E_i are the elementary symmetric functions. Now $|S\Gamma| < 1$. Since $|\Gamma| < 1$ it follows from (11) that $S\Gamma^i = o(S\Gamma)$ for $i > 1$ so, by Newton's identities, $E_2 + \cdots + E_{p-1} = o(S\Gamma)$ and

$$(14) \qquad N\Gamma = -S\Gamma(1 + o(1)).$$

Substituting $B = 1$ in (11) we get $|p| \leqq |\Gamma^{-1} S\Gamma|$, i.e. $|S\Gamma| \geqq |\Gamma p|$. So (14) gives $|N\Gamma| = |\Gamma^p| \geqq |\Gamma p|$ and Theorem 1 follows.

If k has characteristic 0 then it is not possible without completeness assumptions to prove that K/k is generated by a root of an equation of either of the types (1) or (2) (see appendix). The best we can do is to derive, in Theorem 2 below, conditions that there exist an element satisfying a congruence related to (1).

Definition 5. Let $\wp(x)$ denote the polynomial $x^p - x$. If k has characteristic 0 let $\phi(x, y)$ be the polynomial with integal coefficients such that $(x + y)^p = x^p + y^p + p\phi(x, y)$.

Note that if $|A| \geqq |B|$ then $\phi(A, B) = O(A^{p-1}B)$ and that

$$\wp(x + y) = \wp(x) + \wp(y) + p\phi(x, y).$$

THEOREM 2. *Let* K/k *satisfy the assumption of Section* 3. *The following three conditions are equivalent:* K *contains an element* Λ *with*

$$(15) \qquad (\sigma - 1)\Lambda = 1 + o(1);$$

K *contains an element* M *with*

$$(16) \qquad |pM| < 1 \text{ and } SM = -1;$$

$$(17) \qquad |\Gamma^{p-1}| > |p|.$$

If Λ *satisfies* (15) *then* $|\Lambda| \geqq |\Gamma^{-1}|$ *and* $\Lambda = \alpha + \Lambda'$ *where* $\alpha \varepsilon k$, Λ' *satisfies* (15), *and* $|\Lambda'| = |\Gamma^{-1}|$. *For any such* Λ'

$$(18) \qquad (\sigma - 1)\wp(\Lambda') = o(1); \ \wp(\Lambda') = \lambda + o(\Lambda'), \lambda \varepsilon k, |\lambda^{p-1}| < |p^{-p}|.$$

If K/k *is ramified, conditions* (15), (16), (17) *hold if and only if* $|\Gamma| \not\varepsilon |k^{\cdot}|$.

Proof. The first three conditions are true when k has characteristic p so we may assume k has characteristic 0. If Λ satisfies (15) we can write $(\sigma - 1)\Lambda = 1 + p\mathrm{M}$ with $|p\mathrm{M}| < 1$. Then $0 = p + pS\mathrm{M}$ so M satisfies (16). If M satisfies (16) and $\Lambda = \sum_{1}^{p-1} \nu\sigma^{\nu}\mathrm{M}$ then Λ satisfies (15). So the first two conditions are equivalent. From (11) we see that $-\Gamma/S\Gamma$ is an element of minimal value with trace -1 so (16) is solvable if and only if $|p\Gamma| < |S\Gamma|$, which by (14) is equivalent to $|\Gamma^{p-1}| > |p|$. So the three conditions are equivalent.

Now let A be a canonical orthogenerator for K/k. Let Λ satisfy (15) and let $\Lambda = \alpha_0 + \alpha_1 \mathrm{A} + \cdots + \alpha_{p-1}\mathrm{A}^{p-1}$. Then $\Lambda' = \Lambda - \alpha_0$ satisfies (13). Hence $|\Lambda'| = |\Gamma^{-1}(\sigma - 1)\Lambda| = |\Gamma^{-1}|$ and $(\sigma - 1)\Lambda' = 1 + o(1)$. For any such Λ',

$$(\sigma - 1)\wp(\Lambda') = \wp(1 + o(1)) + p\phi(\Lambda', 1 + o(1)) = o(1) + O(p\Lambda'^{p-1}) = o(1)$$

from definition 5 because $|\Lambda'^{p-1}| < |p^{-1}|$ by (17). From (13) we see there is a $\lambda \varepsilon k$ such that $|\wp(\Lambda') - \lambda| = |\Gamma^{-1}(\sigma - 1)\wp(\Lambda')| < |\Gamma^{-1}| = |\Lambda'|$. This proves (18).

Finally assume K/k is ramified and condition (15) holds. Then $|\Lambda'| \not\varepsilon |k^{\cdot}|$ in the previous discussion so $|\Gamma| \not\varepsilon |k^{\cdot}|$. Conversely, if K/k is ramified and $|\Gamma| \not\varepsilon |k^{\cdot}|$ then Γ^{-1} is an orthogenerator and $(\sigma - 1)\delta\Gamma^{-1} = 1 + o(1)$ for some δ in k.

4. Maximally complete fields. Assume that k is maximally complete [2, 4] under its valuation $|\ |$, that K/k is cyclic of degree p with generating automorphism σ, that $|\sigma\mathrm{B}| = |\mathrm{B}|$ for all $\mathrm{B} \varepsilon K$, and that $0 \leqq |p| < 1$. Then ef necessarily equals p, K is also maximally complete, and every pseudo-convergent sequence of elements of K has a pseudo-limit in K. For everything concerning maximally complete fields we follow the definitions of Kaplansky [2] and Schilling [4] except that we use the $|\ |$-notation in place of their v-notation for valuations.

THEOREM 3. *If* K/k *is ramified and the assumptions of Section* 4 *hold*

then $|\Gamma| \varepsilon |k^{\cdot}|$ *if and only if* k *contains a primitive* p*-th root of unity and* $K = k(\beta^{1/p})$ *for a* $\beta \varepsilon k$ *with* $|\beta| \not\varepsilon |k^{\cdot p}|$.

Proof. If $K = k(\beta^{1/p})$ with $|\beta| \not\varepsilon |k^{\cdot p}|$ and k contains primitive p-th roots of unity then K/k is ramified and $\beta^{1/p}$ is an orthogenerator. Since $\sigma(\beta^{1/p}) = \beta^{1/p}\zeta$ for some primitive p-th root of unity ζ, we have $|\Gamma| = |\zeta - 1| \varepsilon |k^{\cdot}|$ from Proposition 2.

Let K/k be ramified and $|\Gamma| \varepsilon |k^{\cdot}|$. By Theorems 1 and 2 $|\Gamma^{p-1}| = |p|$ so k must have characteristic zero. Express Γ by a canonical orthobasis. The values of all nonzero terms are different so we get

$$\Gamma = \gamma + o(\gamma), \quad \gamma \varepsilon k. \tag{19}$$

Then $S\Gamma = S\gamma + S(o(\Gamma)) = p\gamma + o(S\Gamma)$ by (11). Hence $S(\Gamma)(1 + o(1)) = p\gamma$ and $S\Gamma = p\gamma(1 + o(1))$. From (19) we see also that $\sigma^i\Gamma = \gamma + o(\gamma)$ for all i, so $N\Gamma = \gamma^p(1 + o(1))$. By (14) $\gamma^{p-1} = -p(1 + o(1))$. Since k is maximally complete and $|p - 1| = 1$, $1 + o(1)$ is always a $(p-1)$-th power in k [4, p. 61]. So $-p$ is a $(p-1)$-th power in k and since k contains a subfield isomorphic to the p-adic rationals it necessarily [7] contains a primitive p-th root of unity.

Thus $K = k(\mathrm{A})$ with $\mathrm{A}^p = \alpha \varepsilon k$. Since we are assuming $e = p$ and $f = 1$ it is easy to see that we may assume that either $|\alpha| \not\varepsilon |k^{\cdot p}|$ or $\alpha = 1 + o(1)$. But if $\alpha = 1 + o(1)$ then K contains an element M satisfying (16), which by Theorem 2 contradicts the fact that $|\Gamma^{p-1}| = |p|$. Namely, let $\mathrm{A}^p = 1 + o(1) \varepsilon k$. Then $\mathrm{A} = 1 + \Theta$ with $|\Theta| < 1$ and $S\mathrm{A} = 0$ so $S(p^{-1}\Theta) = -1$ and $\mathrm{M} = p^{-1}\Theta$ satisfies (16).

THEOREM 4. *If the assumptions of Section* 4 *hold and* $|\Gamma^{p-1}| > |p|$ *then* K *is generated by a root* Λ *of an equation* (1) *with* $|\Lambda| = |\Gamma^{-1}|$ *and* $(\sigma - 1)\Lambda = 1 + o(1)$.

Proof. Consider the elements Λ of K which satisfy

$$|\Lambda| = |\Gamma^{-1}| \tag{20}$$

and (15). From Theorem 2 it follows that there are such elements. If Λ, Λ' are two of these we put $\Lambda \prec \Lambda'$ whenever

$$|\Lambda' - \Lambda| = |\Gamma^{-1}(\sigma - 1)\wp\Lambda| \tag{21}$$

and

$$|(\sigma - 1)\wp\Lambda'| < |(\sigma - 1)\wp\Lambda|. \tag{22}$$

Let $\mathfrak{P}$ be the family of all well-ordered sets of elements of the type just

described. $\mathfrak{P}$ exists by virtue of Zermelo's *Aussonderungsaxiom,* and is non-empty. For W, $W' \varepsilon \mathfrak{P}$ we shall say W is less than W' when W is an initial segment of W'. Then $\mathfrak{P}$ is partially ordered and every linearly ordered subset of $\mathfrak{P}$ has an upper bound in $\mathfrak{P}$. By the Lemma of Zorn $\mathfrak{P}$ contains a maximal element M.

CONTENTION 1. *M has a last element.*

Proof. From (21) and (22) follows

(23) $\quad |\Lambda'' - \Lambda'| < |\Lambda' - \Lambda|$ whenever $\Lambda \prec \Lambda' \prec \Lambda''$ and $\Lambda, \Lambda', \Lambda'' \varepsilon M$.

So if M has no last element it is a pseudo-convergent sequence. Assume that this is so and let Λ^* be a pseudo-limit of this sequence. Then

(24) $\quad |\Lambda^* - \Lambda| = |\Lambda' - \Lambda|$ whenever $\Lambda \prec \Lambda'$ and $\Lambda, \Lambda' \varepsilon M$.

We shall show that the structure (M, Λ^*) obtained by adjoining Λ^* to M as last term is in $\mathfrak{P}$, contradicting the maximality of M. By (20), (15), and (18) $(\sigma - 1)\wp(\Lambda) = o(1)$ so by (21) $|\Lambda' - \Lambda| < |\Gamma^{-1}|$ and $|\Lambda^*| = |\Gamma^{-1}|$. It remains to check (15) and (22).

Let Λ be any element of M and $\Theta = \Lambda^* - \Lambda$. We have just seen that $\Theta = o(\Gamma^{-1})$. By (12) $(\sigma - 1)\Theta = o(1)$ so Λ^* satisfies (15). Now if $(\sigma - 1)\Theta = \Delta$ then

$$(\sigma - 1)\wp(\Theta) = \wp(\Theta + \Delta) - \wp(\Theta) = \wp(\Delta) + p\phi(\Theta, \Delta) = -\Delta + o(\Delta).$$

We see that (12) implies

(25) $$(\sigma - 1)\wp(\Theta) = -(\sigma - 1)\Theta + o(\Gamma\Theta).$$

This being so, $|\Lambda| = |\Gamma^{-1}|$ and $|\Gamma^{p-1}| > |p|$ imply

(26) $$(\sigma - 1)\wp(\Lambda + \Theta) = (\sigma - 1)\wp(\Lambda) - (\sigma - 1)\Theta + o(\Gamma\Theta).$$

From (21) we see that $|(\sigma - 1)\wp(\Lambda)| = |\Gamma\Theta|$. So from (26)

$$(\sigma - 1)\wp(\Lambda^*) = O(\Gamma\Theta) = O((\sigma - 1)\wp\Lambda).$$

But this holds for every $\Lambda \varepsilon M$ so Λ^* satisfies (22). Thus $(M, \Lambda^*) \varepsilon \mathfrak{P}$ and Contention 1 follows.

CONTENTION 2. *If Λ' is the last element of M then* $(\sigma - 1)\wp(\Lambda') = 0$. *Hence Λ' generates K/k and satisfies an equation* (1).

Proof. Let $(\sigma - 1)\wp(\Lambda') = \Delta$. From Theorem 2 $\Delta = o(1)$. If $\Delta \neq 0$ let Θ be an element of K with $|\Theta| = |\Gamma^{-1}\Delta|$ and $(\sigma - 1)\Theta = \Delta$. Its

existence is guaranteed by (13). Let $\mathrm{A}^* = \mathrm{A}' + \Theta$. We shall show that $(M, \mathrm{A}^*) \varepsilon \mathfrak{P}$ giving a contradiction as before. Everything except the analogue of (22) is clear. For this we apply (26).

$$(\sigma - 1)\wp(\mathrm{A}' + \Theta) = (\sigma - 1)\wp(\mathrm{A}') - (\sigma - 1)\Theta + o(\Gamma\Theta) = o(\Delta).$$

This proves Contention 2 and Theorem 4.

THEOREM 5. *Let K/k satisfy the assumptions of Section 4. If K/k is ramified and $|\Gamma^{p-1}| > |p|$ then $K = k(\mathrm{A})$ where A is an orthogenerator satisfying an equation* (1) *with*

$$(27) \qquad \lambda \notin |k^{\cdot p}|, \quad |\lambda| > 1.$$

If $\bar{K}/\bar{k}$ is inseparable and $|\Gamma^{p-1}| > p|$ then $K = k(\mathrm{A})$ where A is an orthogenerator satisfying an equation (1) *with*

$$(28) \qquad \lambda = \alpha\beta^p, \quad |\alpha| = 1, \quad \bar{\alpha} \notin \bar{k}^{\cdot p}, \quad \beta \varepsilon k, \quad |\beta| > 1.$$

If K/k is separable then $K = k(\mathrm{A})$ where A is an orthogenerator satisfying an equation (1) *with*

$$(29) \qquad |\lambda| = 1, \qquad \bar{\lambda} \notin \wp(\bar{k}^+).$$

Conversely, let k satisfy the assumptions of Section 4. Then every polynomial (1) *either splits completely or has a zero A which generates a cyclic extension of degree p and satisfies* (15) *for some generating automorphism σ. If $K = k(\mathrm{A})$ satisfies the assumptions of Section 4 and λ satisfies* (27), (28), *or* (29) *then K/k is ramified, $\bar{K}/\bar{k}$ is inseparable, or $\bar{K}/\bar{k}$ is separable, respectively, and $|\Gamma^{p-1}| > |p|$.*

Proof. Let $|\Gamma^{p-1}| > |p|$. If K/k is ramified then by Theorems 2 and 4 we have $K = k(\mathrm{A})$ with $|\mathrm{A}| = |\Gamma^{-1}| \notin |k^{\cdot}|$ so $|\lambda| = |\mathrm{A}^p| \notin |k^{\cdot p}|$. If $\bar{K}/\bar{k}$ is inseparable then $|\Gamma| < 1$ and $K = k(\mathrm{A})$ with $|\mathrm{A}| = |\Gamma^{-1}| > 1$ so $|\lambda| = |\mathrm{A}^p| \varepsilon |k^{\cdot p}|$. Let $\lambda = \alpha\beta^p$ with $|\alpha| = 1$, $|\beta| > 1$, $\alpha, \beta \varepsilon k$. Then $|\mathrm{A}\beta^{-1}| = 1$ and $(\mathrm{A}\beta^{-1})^p = \alpha + o(1)$. Suppose $\bar{\alpha} = \bar{\gamma}^p$ with $\gamma \varepsilon k$. Then $\mathrm{A}\beta^{-1} = \gamma + o(1)$, $\mathrm{A} = \beta\gamma + o(\mathrm{A})$, and $(\sigma - 1)\mathrm{A} = o(1)$, which is not true, so $\bar{\alpha} \notin \bar{k}^{\cdot p}$. The case $\bar{K}/\bar{k}$ separable is easily proved.

The proof of Proposition 17 of [7] is easily generalized to prove the statement concerning the behavior of equations (1). If $K = k(\mathrm{A})$ and λ satisfies (27) then K/k is ramified with A an orthogenerator. If λ satisfies (28) then the residue class of $\mathrm{A}\beta^{-1}$ generates an inseparable extension of $\bar{k}$ of degree p. The case when λ satisfies (29) is easily proved. $|\Gamma^{p-1}| > |p|$ was established in Theorem 2.

Warning: There exist irreducible equations (1) which do not satisfy

(27), (28), or (29). If $|\lambda^{p-1}| > |p^{-p}|$ they sometimes generate cyclic extensions with $|\Gamma^{p-1}| > |p|$ but the value of Γ cannot be read off directly from the value of λ.

Definition 6. A field generated by a root of an equation (1) is called an *Artin-Schreier extension.*

PROPOSITION 3. *If $\bar{k}$ has no inseparable extensions then every cyclic subfield of a composite of Artin-Schreier extensions of k is an Artin-Schreier extension.*

Proof. By Theorem 4 we can assume k contains primitive p-th roots of unity. Then the Artin-Schreier extensions are those obtained by adjoining a p-th root of an element of value 1, by the proof of Theorem 3.

We do not know whether the assumption on $\bar{k}$ can be omitted.

5. Explicit reciprocity law. Let k satisfy the assumptions of Section 4. Let C be the algebraic closure of k and G the Galois group of C/k [1]. Let $\mathfrak{l}$ be the set of all $\lambda \varepsilon k$ such that $|\lambda^{p-1}p^p| < 1$.

If $\lambda \varepsilon \mathfrak{l}$ and $\sigma \varepsilon G$ choose a $\Lambda \varepsilon C$ such that $\wp(\Lambda) = \lambda$ and denote by $\lambda^*\sigma$ the residue class modulo p of an integer ν such that $(\sigma - 1)\Lambda = \nu + o(1)$. ν exists by Theorem 5. It is easy to verify that $\lambda^*\sigma$ does not depend upon the choice of Λ and that $\lambda^*\sigma\tau = \lambda^*\sigma + \lambda^*\tau$ if $\sigma, \tau \varepsilon G$. Hence λ^* is a character of G of period p.

The mapping $\lambda \rightarrow \lambda^*$ is not in general a homomorphism (as it is when k has characteristic p). Namely, if $|p|$ is small enough it can happen that both λ and $p\lambda$ lead to equations of type (28). Then $(p\lambda)^* \neq 0$ but $p(\lambda^*) = 0$. We shall now develop a sufficient condition that $(\lambda + \mu)^* = \lambda^* + \mu^*$.

PROPOSITION 4. *If $\Lambda \varepsilon C$, $\lambda \varepsilon \mathfrak{l}$, and $\wp(\Lambda) = \lambda + o(1)$ then $k(\Lambda)$ contains a Λ_0 with $\Lambda_0 = \Lambda + o(1)$ and $\wp(\Lambda_0) = \lambda$.*

Proof. If $\Lambda' \varepsilon k(\Lambda)$ and $\wp(\Lambda') = \lambda + o(1)$ let $\Delta' = \wp(\Lambda') - \lambda$ and define $\Lambda'' = \Lambda' + \Delta'$. Then $\wp(\Lambda'') = \wp(\Lambda') + \wp(\Delta') + p\phi(\Lambda', \Delta')$. By definition of $\mathfrak{l}$, $|p\Lambda'^{p-1}| < 1$; so $\wp(\Lambda'') = \lambda + o(\Delta')$. Our proposition follows by applying the Lemma of Zorn just as in the proof of Theorem 4. Of course, the Lemma of Zorn can be avoided when the valuation is discrete and of rank 1.

COROLLARY. *If $|\lambda| < 1$ then $\lambda^* = 0$.*

PROPOSITION 5. *If* $|\mu| \leqq |\lambda|$, $\lambda \in \mathfrak{l}$, *and* $|\lambda^{p-1}\mu p^p| < 1$ *then* $\mu^* + \lambda^* = (\mu + \lambda)^*$.

Proof. Let $\wp(\Lambda) = \lambda$, $\wp(\mathrm{M}) = \mu$. Let $\mathrm{N} = \Lambda + \mathrm{M}$. Then

$$\wp(\mathrm{N}) = \lambda + \mu + p\phi(\Lambda, \mathrm{M}) = \lambda + \mu + o(1).$$

By Proposition 4 there is an N_0 in $k(\mathrm{N})$ such that $\mathrm{N}_0 = \mathrm{N} + o(1)$ and $\wp \mathrm{N}_0 = \lambda + \mu$. If $\sigma \in G$ then

$$(\sigma - 1)\mathrm{N}_0 = (\sigma - 1)\mathrm{N} + o(1) = \phi^*\sigma + \mu^*\sigma + o(1)$$

which was to be proved.

In particular, if $|\mu| = 1$ and $\lambda \in \mathfrak{l}$ then we have $\mu^* + \lambda^* = (\mu + \lambda)^*$.

For each $\lambda \in \mathfrak{l}$ we can form the 2-cocycle (λ^*, α, k) according to [7] or the algebra $(\alpha, \lambda]$ according to [10], which amounts to the same thing. Consider the symbol $(\alpha, \lambda]$ as denoting the corresponding 2-cohomology class as in [7]. Using Theorem 1 of [7] and Proposition 5 of this paper we obtain the following rules:

1. *For all* $\lambda \in \mathfrak{l}$, $(\lambda, -\lambda] = 1$.
2. *If* $\lambda, \mu \in \mathfrak{l}$, $|\mu| \leqq |\lambda|$, *and* $|\lambda^{p-1}\mu p^p| < 1$ *then* $(\alpha, \lambda]\,(\alpha, \mu] = (\alpha, \lambda + \mu]$. *This is always true if* $\lambda \in \mathfrak{l}$ *and* $|\mu| = 1$.
3. $(\alpha, \lambda]\,(\beta, \lambda] = (\alpha\beta, \lambda]$.

Rule 1 follows from the fact that $-\lambda$ is the norm of $-\Lambda$ if $\wp(\Lambda) = \lambda$.

We now assume that k is a regular field of generalized local class field theory [7].

Let $\lambda = \lambda'\pi^{-c} \in \mathfrak{l}$, $|\lambda'| = 1$, and $p \nmid c$. We wish to compute the invariant of the algebra $(\beta, \lambda]$. If we could do this for every $\beta \in k^{\cdot}$ we would have a complete explicit reciprocity law. As it is, we are able to compute the invariant of $(\beta, \lambda]$ for a special class of elements β which form a basis for $k^{\cdot}$ modulo $N(K/k)$.

Let α be an element of k such that $K(\mathrm{A})/k$ is unramified if $\wp(\mathrm{A}) = \alpha$ and such that the invariant $\rho(\pi, \alpha]$ (see [7]) is $[1/p]$ (mod 1). Then, since $|\alpha - \lambda| = |\pi^{-c}|$, $\rho(\alpha - \lambda, \alpha] = [-c/p]$ (mod 1). By rule 1,

$$(\alpha - \lambda, -(\alpha - \lambda]\;(\alpha - \lambda, \alpha] = (\alpha - \lambda, \alpha].$$

By rule 2,

$$(\alpha - \lambda, -(\alpha - \lambda)]\,(\alpha - \lambda, \alpha] = (\alpha - \lambda, \lambda].$$

Hence $\rho(\alpha - \lambda, \lambda] = \rho(\alpha - \lambda, \alpha] = [-c/p]$ (mod 1). If one prefers a β

of value 1 instead of $|\pi^{-c}|$ then $\rho(1-\alpha\lambda^{-1},\lambda]=[-c/p]\ (\mathrm{mod}\,1)$ since $-\lambda^{-1}$ is a norm. Thus we obtain the law

$$\rho(1-\alpha\lambda^{-1},\lambda]=[-c/p]\ (\mathrm{mod}\,1). \tag{30}$$

We are now in a position to prove the explicit reciprocity law for the field of primitive p^2-th roots of unity over the p-adic completion of the rational field. If this problem were solved for the field of p^n-th roots of unity it would make possible an elegant proof of Artin's global reciprocity law, for it is well known that this law can be proved from the global index theorems and the global reciprocity law for cyclotomic fields.

Appendix.

SCHOLIUM 1. *It is not possible without assumptions on k in addition to those of Sections* 1, 2, 3 *to prove that every cyclic K/k with $|\Gamma^{p-1}|>|p|$ is generated by a root of an equation* (1).

Proof. Let r denote the rational field, $|\ |$ the p-adic valuation, and C the cyclic subfield of degree p of the field of primitive p^2-th roots of unity. Then $e=p$. If r_p denotes the p-adic rationals then Cr_p/r_p is generated by a root of $x^p-x-p^{-1}=0$ (see [8]) so $|\Gamma^{p-1}|>|p|$. But for p odd C/r is cyclic of odd degree so any defining equation must split into real linear factors in C. By Descartes' rule of signs an equation $x^p-x-a=0$, for a real, can have at most three real roots. Hence C/r is not generated by a root of any such equation for $p>3$.

Unfortunately this method does not disprove the conjecture: if $|\Gamma^{p-1}|>|p|$ and k contains primitive $(p-1)$-th roots of unity then K is generated by a root of an equation (1). Although this seems very unlikely we have not been able to find a counterexample.

SCHOLIUM 2. *Without assumptions in addition to those of Sections* 1, 2, 3 *it is not possible to prove that every irreducible equation* (1) *generates a cyclic extension for $p>2$.*

Proof. Consider x^p-x-1 over the rational field. Being irreducible modulo p is it irreducible. For $p>3$ Descartes' rule of signs shows that its splitting field has even degree and this is easy to check for $p=3$ also.

SCHOLIUM 3. *If k is not complete then K/k cyclic, $e=p$, $|\Gamma^{p-1}|=|p|$ can happen without primitive p-th roots of unity being contained in k.*

Proof. $r(\sqrt{6})$ does not contain primitive cube roots of unity (being

real) but $r_3(\sqrt{6})$ does. So by Wang's Theorem [5] it is easy to see that there is a cyclic cubic extension of $r(\sqrt{6})$ which is a counterexample.

SCHOLIUM 4. *If $\bar{k}$ has an inseparable extension there may exist a K/k which is cyclic of degree p, does not contain primitive p-th roots of unity, and contains no Λ with $(\sigma-1)\Lambda = 1 + o(1)$.*

Proof. Let p be any odd prime. Let k_0 be the completion of the field of all rational functions with rational coefficients of a transcendental element t, under the valuation which defines the value of a polynomial in t to be the maximum p-adic value of its coefficients. Let $k = k_0(((1-t^{p-1})p)^{1/(p-1)})$. Let K be the splitting field over k of the polynomial $x^p - px - t$.

INDIANA UNIVERSITY AND INSTITUTE FOR ADVANCED STUDY.

REFERENCES.

[1] E. Artin, *Algebraic Numbers and Algebraic Functions. I.* Mimeographed lecture notes, Princeton University, New York University, 1951.

[2] I. Kaplansky, "Maximal fields with valuation," *Duke Mathematical Journal*, vol. 9 (1942), pp. 303-321.

[3] O. Ore, "Abriss einer arithmetischen Theorie der Galoisschen Körper," Parts 1 and 2, *Mathematische Annalen*, vol. 100 (1928), pp. 650-673, and vol. 102 (1930), pp. 283-304.

[4] O. F. G. Schilling, *The Theory of Valuations*, Mathematical Surveys, no. 4, American Mathematical Society, New York, 1950.

[5] Shianghaw Wang, "On Grunwald's theorem," *Annals of Mathematics*, vol. 51 (1950), pp. 471-484.

[6] G. Whaples, "Existence of generalized local class fields," *Proceedings of the National Academy of Sciences*, vol. 39 (1953), pp. 1100-1103.

[7] ———, "Generalized local class field theory, I and II," *Duke Mathematical Journal*, vol. 19 (1952), pp. 505-517, and vol. 21 (1954), pp. 247-255.

[8] ———, "Generalized local class field theory, IV, Cardinalities," *ibid.*, vol. 21 (1954), pp. 583-586.

[9] E. Witt, "Der Existenzsatz für abelsche Funktionenkörper," *Journal für die reine und angewandte Mathematik*, vol. 173 (1935), pp. 43-51.

[10] ———, "Schiefkörper über diskret bewerteten Körpern," *ibid.*, vol. 176 (1936), pp. 153-156.

REMARK ON AN APPLICATION OF PSEUDOANALYTIC FUNCTIONS.*[1]

By LIPMAN BERS.

1. Introduction. In this note we show how the theory of pseudoanalytic functions as formulated in [5] yields very precise information on the behaviour of solutions of a linear ellipitic equation

$$L\phi = a_{11}\phi_{xx} + 2a_{12}\phi_{xy} + a_{22}\phi_{yy} + a_1\phi_x + a_2\phi_y + a_0\phi = 0 \tag{1.1}$$

near a regular or isolated singular point. The theorems stated below (Section 2) contain and are stronger than the previous results on local behaviour of solutions of (1.1) due to the author [1, 2, 4], Vekua [16] and Hartman and Wintner [10, 12]. These theorems refer to single-valued solutions. But the same method would give analogous results for finitely-many-valued solutions and for solutions with finitely-many-valued gradients.

Without loss of generality we consider equations and solutions defined near the origin of the z-plane $(z = x + iy)$ and assume that

$$a_{11} = a_{22} = 1, \qquad a_{12} = 0 \text{ at } z = 0. \tag{1.2}$$

Concerning the smoothness of the coefficients in (1.1) we make either of the following assumptions.

HYPOTHESIS α. *The leading coefficients $a_{ij}(x,y)$ of (1.1) satisy a uniform Hölder condition and condition (1.2). The coefficients $a_i(x,y)$ are measurable functions which belong to the space L_p for some $p > 2$.*

HYPOTHESIS β. *All coefficients of (1.1) satisfy a uniform Hölder condition and (1.2) holds.*

Under Hypothesis β we require that solutions of (1.1) be of class C^2. Under Hypothesis α we have no right to expect twice continuously differen-

* Received November 28, 1955.

[1] Work supported by the Office of Ordnance Research, United States Army, under Contract No. DA-30-069-ORD-835. Reproduction in whole or in part is permitted for any purpose of the United States Government.

tiable solutions. We shall say that ϕ is a solution if the derivatives ϕ_x, ϕ_y exist, are continuous and possess generalized L_p derivatives $\phi_{xx}=(\phi_x)_x$, $\phi_{xy}=(\phi_x)_y=(\phi_y)_x$ and $\phi_{yy}=(\phi_y)_y$ which satisfy (1.1) almost everywhere. The equality of the mixed derivatives is a consequence of the definition of generalized derivatives.

In Section 3 we shall show, using an inequality of Calderón and Zygmund [7], that equation (1.1) can be reduced to an equation involving only second derivatives (Theorem I), to a system of two first order equations (Theorem II), and to the equations characterizing pseudoanalytic functions (Theorem III). This will yield the desired description of the local behaviour of solutions (Theorems A and B of Section 2).

Hypothesis α is essentially sharp, since our results are certainly not true for equations with continuous but not Hölder continuous coefficients (cf. Hartman and Wintner [12]). Concerning the local behaviour of solutions of elliptic equations with discontinuous coefficients we refer to a recent paper by Nirenberg and the author [6].

For the convenience of the reader we recall the definition of generalized derivatives (Friedrichs [8], Sobolev [15]). Let $f(x,y)$, $g(x,y)$, $h(x,y)$ be measurable functions defined in a domain D. If $|f|^p$, $|g|^p$, $|h|^p$ are locally summable ($p\geqq 1$), the statement

$$f_x=g,\ f_y=h \text{ in the } L_p\text{-sense}$$

means that

(a) in every compact subset D_0 of D there exists a sequence of continuously differentiable functions $f^{(n)}$ such that

$$\int\int_{D_0}\{|f^{(n)}-f|^p+|f_x^{(n)}-g|^p+|f_y^{(n)}-h|^p\}dxdy\rightarrow 0,$$

(b) for every continuously differentiable function ω which vanishes identically near the boundary of D

$$\int\int_D f\omega_x\,dxdy=-\int\int_D g\omega\,dxdy,\quad \int\int_D f\omega_y\,dxdy=-\int\int_D h\omega\,dxdy,$$

and

(c) for almost every y and for almost every x

$$\int_\alpha^\beta g(\xi,y)d\xi=f(\beta,y)-f(\alpha,y),\qquad \int_\gamma^\delta h(x,\eta)d\eta=f(x,\delta)-f(x,\gamma)$$

for almost all $\alpha,\beta,\gamma,\delta$.

Each of the three conditions (a), (b), (c) implies the other two. It is known that f is Hölder continuous if f_x and f_y exist in the L_p-sense for some $p > 2$.

2. Statement of results. Let $\phi(z) = \phi(x, y)$ be a solution of (1.1) defined in a neighborhood of the origin. Assuming Hypothesis α we say that ϕ has a zero of integral order $n > 0$ at the origin if for some complex $\alpha \neq 0$

$$(2.1) \qquad \phi \sim \operatorname{Re}(\alpha z^n), \qquad \phi_x - i\phi_y \sim n\alpha z^{n-1}.$$

In the case of Hypothesis β we also require that

$$(2.2) \qquad \phi_{xx} - i\phi_{xy} \sim -\phi_{yy} - i\phi_{xy} \sim n(n-1)\alpha z^{n-2}.$$

Here and hereafter all asymptotic relations refer to $z \to 0$.

THEOREM A. *Under Hypotheses α or β a solution of* (1.1) *which vanishes at the origin without vanishing identically has at the origin a zero of integral order.*

The following consequence of Theorem A is of interest for differential geometry in the large (cf. Hartman and Wintner [11] and the references given there).

COROLLARY. *Let the function $\Omega(x, y, \phi, p, q, r, s, t)$ be a Hölder continuously differentiable function of its eight variables, defined for sufficiently small values of the variables, and satisfying the condition: $4\Omega_r\Omega_t - \Omega_s^2 > 0$ at $x = y = \phi = p = q = r = s = t = 0$. Let $\phi'(x, y)$ and $\phi''(x, y)$ be two twice continuously differntiable functions defined in the neighborhood of the origin and satisfying the differential equation*

$$(2.3) \qquad \Omega(x, y, \phi, \phi_x, \phi_y, \phi_{xx}, \phi_{xy}, \phi_{yy}) = 0.$$

Assume that the functions ϕ', ϕ'' vanish at the origin together with their derivatives of the first and second order, but the difference $\phi = \phi'' - \phi'$ is not identically zero. Then the expression $\phi_{xx}\phi_{yy} - \phi_{xy}^2$ is negative in a deleted neighborhood of the origin.

The conclusion of the Corollary was stated, proved and applied to the uniqueness proof for Minkowski's problem by H. Lewy [13], under the assumption that the function Ω, and hence also every solution of the differential equation (2), is analytic. Recently Hartman and Wintner [9, 10, 11] extended it to the case when Ω is twice continuously differentiable. The

present formulation seems to be essentially sharp. It will be seen from the proof, however, that for special cases (for instance, for quasi-linear equations) weaker conditions are sufficient.

In order to prove the Corollary observe that the difference $\phi = \phi'' - \phi'$ satisfies a linear elliptic partial differential equation of the form (1.1) with

$$a_{11}(x, y) = \int_0^1 \Omega_r[x, y, (1-\lambda)\phi'(x, y) + \lambda\phi''(x, y), \cdots]d\lambda$$

and the other coefficients a_{ij}, a_i defined similarly. We may assume that (1.2) holds since this can be achieved by an an affine transformation. According to a theorem of Nirenberg [14] the second derivatives of the solutions ϕ', ϕ'' satisfy a Hölder condition. Hence the coefficients a_{ij}, a_i satisfy such a condition, and Theorem A implies that $\phi_{xx}\phi_{yy} - \phi_{xy}{}^2 \sim - n^2(n+1)^2 \mid \alpha \mid^2 \mid z \mid^{2n-2}$ with $\mid \alpha \mid > 0$ and $n > 1$.

We consider next a single-valued solution ϕ of (1.1) defined in a deleted neighborhood of the origin. This solution will be said to have at the origin a logarithmic singularity if (under Hypothesis α) for some real $\alpha \neq 0$

$$(2.4) \qquad \phi \sim \alpha \log \mid z \mid, \qquad \phi_x - i\phi_y \sim \alpha/z$$

and (under Hypothesis β) also

$$(2.5) \qquad \phi_{xx} - i\phi_{xy} \sim -\phi_{yy} - i\phi_{xy} \sim -\alpha/z^2.$$

We say that the origin is a pole of ϕ if (under Hypothesis α)

$$(2.6) \qquad \phi \sim \operatorname{Re}(\alpha z^{-n}), \qquad \phi_x - i\phi_y \sim -n\alpha z^{-n-1}$$

for some complex $\alpha \neq 0$ and integer $n > 0$, and (under Hypothesis β) also

$$(2.7) \qquad \phi_{xx} - i\phi_{xy} \sim -\phi_{yy} - i\phi_{xy} \sim n(n+1)z^{-n-2}.$$

Finally, $z = 0$ will be called an essential singularity of ϕ if (under Hypothesis α)

$$(2.8) \qquad \limsup (\mid z \mid^N \phi) = \limsup (-\mid z \mid^N \phi) = \limsup (\mid z \mid^N \mid \phi_x - i\phi_y \mid) = +\infty$$

and (under Hypothesis β) also

$$(2.9) \quad \limsup (\mid z \mid^N \mid \phi_{xx} - i\phi_{xy} \mid) = \limsup (\mid z \mid^N \mid \phi_{yy} + i\phi_{xy} \mid) = +\infty$$

for every $N > 0$. If $a_0 \equiv 0$, we require also that

$$(2.10) \qquad \liminf \mid \phi_x - i\phi_y - \gamma \mid = 0$$

for every complex γ.

THEOREM B. *Under Hypotheses α or β let ϕ be a single-valued solution of* (1.1) *defined in a deleted neighborhood of the origin. Then the singularity of ϕ at $z=0$ is either removable, or logarithmic, or a pole, or essential.*

The theorem implies, in particular, that if ϕ' and ϕ'' are two solutions having at $z=0$ logarithmic singularities, then there exist two real constants λ_1, λ_2 such that $\lambda_1\phi'+\lambda_2\phi''$ is regular at the origin.

3. Reduction to pseudoanalytic functions. This reduction will be accomplished in three steps.

We remark that in the proofs of Lemma 3.1 and Theorem I given below the Hölder continuity of the leading coefficients and the fact that the number of independent variables is two are not used, and the conclusions would remain true, with obvious modifications, in a more general case, say for continuous a_{ik}.

LEMMA 3.1. *Under Hypothesis α there exists a solution ϕ of equation* (1.1) *which is defined in a neighborhood of the origin and satisfies the conditions*

$$\phi=c_0,\ \phi_x=c_1,\ \phi_y=c_2 \text{ at } z=0, \tag{3.1}$$

where c_0, c_1, c_2 are given numbers.

The proof is a variant of the classical Korn argument based on the results of Calderón and Zygmund [7]. Let $\boldsymbol{B}$ denote the Banach space of real functions $\phi(z)$ defined for $|z|\leqq R$ and possessing continuous first and generalized second derivatives, for which the norm

$$\|\phi\|=\max(|\phi|,|\phi_x|,|\phi_y|)+\{\int\int_{|z|<R}(|\phi_{xx}|+|\phi_{yy}|+|\phi_{yy}|)^p\,dxdy\}^{1/p}$$

is finite. Define the mapping $\boldsymbol{T}$ of $\boldsymbol{B}$ into itself by defining $\chi=\boldsymbol{T}\phi$ to be the function

$$\chi(z)=\frac{1}{2\pi}\int\int_{|\zeta|<R}[(\Delta-\boldsymbol{L})\phi(\zeta)]\log|\zeta-z|\,d\xi d\eta.$$

Using the inequalities of Hölder and of Calderón-Zygmund it is easy to show that $\boldsymbol{T}$ is bounded, that $\Delta\boldsymbol{T}=\Delta-\boldsymbol{L}$, and that $\|\boldsymbol{T}\|\to 0$ for $R\to 0$. If $h(z)$ is a harmonic function with $\|h\|<+\infty$, and R is so small that $\|\boldsymbol{T}\|\leqq\theta<1$, then the equation

$$\phi=\boldsymbol{T}\phi+h \tag{3.2}$$

has a solution which satisfies equation (1.1) and the inequality

$$\| \phi - h \| \leqq \theta(1-\theta)^{-1} \| h \|.$$

We solve equation (3.2) for a sufficiently small R and for $h = 1, x, y$, and obtain three solutions ϕ_0, ϕ_1, ϕ_2 of (1.1) for which the values of

$$| \phi_0 - 1 |, | \phi_{0,x} |, | \phi_{0,y} |, | \phi_1 |, | \phi_{1,x} - 1 |, \ | \phi_{1,y} |, | \phi_2 |, | \phi_{2,x} | \ | \phi_{2,y} - 1 |$$

are very small. A properly chosen linear combination of these solutions satisfies (3.1).

LEMMA 2.2. *Under Hypothesis β there exists a solution ϕ of equation (1.1) which is defined in the neighborhood of the origin and satisfies the conditions*

$$(2.3) \qquad \phi = c_0, \ \phi_x = c_1, \ \phi_y = c_2, \ \phi_{xx} = -\phi_{yy} = c_{11}, \ \phi_{xy} = c_{12} \ \text{at } z = 0,$$

where c_0, c_1, c_2, c_{11}, c_{12} are given numbers.

This (known) result is proved by considering $\boldsymbol{T}$ as an operator in the Banach space of functions ϕ defined in $| z | \leqq R$ and having second derivatives satisfying a uniform Hölder condition with exponent $\epsilon < 1$, ϵ being a Hölder exponent for the coefficients of (1.1). The norm in this space is

$$\| \phi \| = \max(| \phi |, R\, | \phi_x |, R\, | \phi_y |, \ R^2\, | \phi_{xx} |, R^2\, | \phi_{xy} |, R^2\, | \phi_{yy} |) + R^{2+\epsilon} K,$$

where K is the smallest Hölder constant for ϕ_{xx}, ϕ_{xy}, ϕ_{yy} belonging to the exponent ϵ. The desired function is obtained by solving equation (3.2) for a small R and for $h = 1, x, y, x^2 - y^2, xy$.

THEOREM I. *Under Hypothesis α there exist, in a neighborhood of the origin, three functions ϕ_0, ξ, η such that*

$$\boldsymbol{L}\phi_0 = \boldsymbol{L}\xi = \boldsymbol{L}\eta = 0,$$

$$(3.4) \qquad \phi_0 = 1, \ \phi_{0,x} = \phi_{0,y} = 0 \ \text{at } z = 0,$$

$$(3.5) \qquad \xi = \eta = \xi_y = \eta_x = 0, \ \xi_x = \eta_y = 1 \ \text{at } z = 0.$$

A function $\phi(x, y)$ defined in a neighborhood of the origin is a solution of (1.1) if and only if the function $\Phi = \phi/\phi_0$, considered as a function of (ξ, η) is a (twice continuously differentiable) solution of the equation

$$(3.6) \qquad A_{11}\Phi_{\xi\xi} + 2A_{12}\Phi_{\xi\eta} + A_{22}\Phi_{\eta\eta} = 0$$

where the A_{ij} are certain functions satisfying a Hölder condition and the condition

$$A_{11} = A_{12} = 1, \; A_{12} = 0 \text{ at } \zeta = \xi + i\eta = 0. \tag{3.7}$$

Under Hypothesis β the same conclusion holds and the functions ϕ, ξ, η may be chosen so that

$$\phi_{0,xx} = \phi_{0,xy} = \phi_{0,yy} = \xi_{xx} = \xi_{xy} = \xi_{yy} = \eta_{xx} = \eta_{xy} = \eta_{yy} = 0 \text{ at } z = 0. \tag{3.8}$$

Proof. The existence of solutions of (1.1) satisfying conditions (3.4), (3.5), (3.8) follows from Lemmas 3.1 and 3.2. The equivalence of equations (1.1) and (3.6) is proved by a direct computation. In the case of Hypothesis α the legitimacy of this computation may be established by a simple argument.

The Hölder continuity of the functions A_{ij} follows from the Hölder continuity of the first derivatives of solutions of (1.1) under Hypothesis α (cf. Section 1).

In order to complete the proof we must show that every solution of (3.6) which has continuous first derivatives and generalized second derivatives in L_p $(p \geqq 2)$ also has continuous second derivatives. This, however, is an immediate consequence of two statements: the Dirichlet problem for an elliptic equation (3.6) with Hölder continuous coefficients has a unique twice continuously differentiable solution; every solution of (3.6) with continuous first and generalized second L_p-derivatives, $p \geqq 2$, obeys the maximum principle. The first statement is classical; the second has been established recently (Bers and Nirenberg [6]).

THEOREM II. *Let there be given an elliptic equation*

$$a_{11}\phi_{xx} + 2a_{12}\phi_{xy} + a_{22}\phi_{yy} = 0 \tag{3.9}$$

satisfying Hypothesis β. There exist four Hölder continuous functions $b_{11}, \cdots, b_{22}$ defined in a neighborhood of the origin, satisfying the condition

$$b_{11} = b_{22} = 0, \; b_{12} = -b_{21} = 1 \text{ at } z = 0 \tag{3.10}$$

and such that in a neighborhood of the origin equation (3.9) is equivalent to the elliptic system

$$\phi_x = b_{11}\psi_x + b_{12}\psi_y, \qquad \phi_y = b_{21}\psi_x + b_{22}\psi_y. \tag{3.11}$$

This means that to every solution ϕ of (3.9) there exists a (not necessarily single-valued) function ψ satisfying (3.11), and that whenever ϕ and

ψ have continuous derivatives and satisfy (3.11), ϕ is a twice continuously differentiable solution of (3.9).

Addition to Theorem II. *Let the coefficients of* (3.9) *be Hölder continuous and satisfy the ellipticity condition* $(a_{11}a_{22}-a_{12}^2>0)$ *in a domain* D. *Then there exist Hölder continuous functions* b_{ij} *defined in the whole domain* D, *satisfying* (3.10), *and such that system* (3.11) *is elliptic* $(-4b_{12}b_{21}-b_{11}^2-b_{22}^2+2b_{11}b_{22}>0)$ *and equivalent to* (3.9).

Proof. Consider the elliptic system

$$\lambda_x=(2a_{12}/a_{22})\mu_x+\mu_y, \qquad \lambda_y=-(a_{11}/a_{22})\mu_x. \tag{3.12}$$

Using the Korn-Lichtenstein method (cf. the proof of Lemmas 2.1, 2.2) or the theorem on univalent solutions of elliptic systems [3] one obtains easily a solution (λ,μ) with Hölder continuous derivatives, defined in the neighborhood of the origin and satisfying the conditions

$$\lambda_x=\mu_y=0, \quad -\lambda_y=\mu_x=1 \text{ at } z=0. \tag{3.13}$$

With these functions form the system

$$\chi_x=\lambda_x\phi_x+\mu_x\phi_y, \qquad \chi_y=\lambda_y\phi_x+\mu_y\phi_y. \tag{3.14}$$

We claim that near $z=0$ every solution ϕ of (3.9) is a solution of (3.14), that is that

$$I_C=\int_C(\lambda_x\phi_x+\mu_x\phi_y)\,dx+(\lambda_y\phi_x+\mu_y\phi_y)\,dy=0,$$

where C is a simple closed smooth curve located near the origin and homotopic to zero in the domain of definition of ϕ. This would follow at once from Green's theorem and (3.12), if the functions λ,μ were of class C^2. Under the present circumstances the proof can be accomplished by approximating λ and μ by C^2 functions.

Conversely, if (χ,ϕ) satisfies (3.14), ϕ is a solution of (3.9). This follows from the unique solvability of the Dirichlet problem for (3.9) since solutions of (3.14) obey the maximum principle (cf., for instance, [6]).

Solving system (3.14) algebraically for ϕ_x and ϕ_y and setting $\psi=-\chi$ we obtain the desired system (3.11).

In order to prove the Addition to Theorem II we need a solution of (3.12) defined and satisfying the condition $\mu_x>0$ in the whole domain D. Such a solution can be obtained by the method used in [3]. We give no details since the in-the-large result will not be used in the sequel.

THEOREM III. *Under the hypothesis of Theorem* II *let* $\zeta=\xi(x,y)$

$+i\eta(x,y)$ be a Hölder continuously differentiable homeomorphism of a neighborhood of the origin, with

$$\xi=\eta=\xi_y=\eta_x=0,\ \xi_x=\eta_y=1 \text{ at } z=0, \tag{3.15}$$

which is conformal with respect to the metric

$$a_{22}dx^2-2a_{12}\,dxdy+a_{22}dy^2. \tag{3.16}$$

Set

$$G(\zeta)=-(b_{11}+b_{22})/2+i[-b_{12}b_{21}-(b_{11}-b_{22})^2/4]^{\frac{1}{2}},$$

$$\Gamma(\zeta)=-a_{12}/a_{11}+i(a_{11}a_{22}-a_{12})^{\frac{1}{2}}/a_{11}.$$

Let ϕ be a solution of (3.9) defined in a (perhaps deleted) neighborhood of the origin, ψ the function connected with ϕ by equations (3.11), and set

$$\omega(\zeta)=\phi+i\psi,\qquad w(\zeta)=\phi+G\psi,$$

$$\Omega(\zeta)=\phi_x-i\phi_y,\qquad W(\zeta)=\phi_x-\Gamma\phi_y.$$

Then ω and Ω are pseudoanalytic functions (of the second kind) with generators $(1,G)$ and $(1,\Gamma)$, respectively, w and W are the corresponding pseudoanalytic functions of the first kind, and

$$\omega\sim w, \tag{3.17}$$ $$\Omega\sim W\sim\dot{w}, \tag{3.18}$$

$$\phi_{xx}-i\phi_{xy}\sim-\phi_{yy}+i\phi_{xy}\sim\dot{W}, \tag{3.19}$$

where $\dot{w}$ is the $(1,G)$ derivative of $w(\zeta)$ and W the $(1,\Gamma)$ derivative of $\dot{W}(\zeta)$.

Proof. The existence of the homeomorphism ζ is a classical result of Lichtenstein. Conformality with respect to (3.10) means, of course that

$$\xi_x{}^2+\eta_x{}^2=\nu a_{22},\ \xi_x\xi_y+\eta_x\eta_y=-\nu a_{12},\ \xi_y{}^2+\eta_y{}^2=\nu a_{11},$$

where $\nu>0$. A direct computation shows that the mapping $z\rightarrow\zeta$ takes system (3.11) into the system

$$\phi_\xi=\tau\psi_\xi+\sigma\psi_\eta,\qquad \phi_\eta=-\sigma\psi_\xi+\tau\psi_\eta, \tag{3.20}$$

where

$$2\tau=b_{11}+b_{22},\ \sigma^2+\tau^2=b_{11}b_{22}-b_{12}b_{21},\ \sigma>0.$$

These equations express the $(1,G)$ pseudoanalyticity of ω and w. Since $G(0)=i$ in view of (3.10), relation (3.17) follows.

Next, set $\Phi=\phi_x$, $\Psi=-\phi_y$. By (3.9)

$$\Phi_x=(2a_{12}/a_{11})\Psi_x+(a_{22}/a_{11})\Psi_y,\qquad \Phi_y=-\Psi_y.$$

The mapping $z \to \zeta$ takes this system into the system

$$\Phi_\xi = \tau^* \Psi_\xi + \sigma^* \Phi_\eta, \qquad \Phi_\eta = -\sigma^* \Psi_\xi + \tau^* \Psi_\eta, \tag{3.21}$$

where

$$2\tau^* = 2a_{12}/a_{11}, \quad \sigma^{*2} + \tau^{*2} = a_{22}/a_{11}, \quad \sigma^* > 0.$$

These equations express the $(1, \Gamma)$ pseudoanalyticity of $\Omega = \Phi + i\Psi$ and $W = \Phi + \Gamma\Psi$. Since $\Gamma(0) = i$ in view of (1.2), $\Omega \sim W$.

By the definition of the derivative of a pseudoanalytic function,

$$2\dot{w} = \phi_\xi - i\phi_\eta + (\psi_\xi - i\psi_\eta) G, \text{ and } \phi_\xi + i\phi_\eta + (\psi_\xi + i\psi_\eta) G = 0,$$

by (3.20); so that $\dot{w} \sim \phi_\xi - i\phi_\eta$. But by (3.15), $\Omega = \phi_x - i\phi_y \sim \phi_\xi - i\phi_\eta$, so that (3.18) follows.

Also,

$$2\dot{W} = (\Phi_\xi - i\Phi_\eta) + (\Psi_\xi - i\Psi_\eta)\Gamma, \text{ and } \Phi_\xi + i\Phi_\eta + (\Psi_\xi + i\Psi_\eta)\Gamma = 0,$$

by (3.21); so that $\Phi_\xi - i\Phi_\eta \sim \dot{W}$. But by (3.18), $\Phi_\xi \sim \Phi_x = \phi_{xx}$ and $\Phi_\eta \sim \Phi_y = -\phi_{xy}$ and by (1.2), (3.9) $\phi_{xx} - i\phi_{xy} \sim -\phi_{yy} - i\phi_{xy}$, so that (3.19) is proved.

4. Proof of Theorems A and B. In view of Theorem I it suffices to prove Theorems A and B for an equation of the form (3.6) with Hölder continuous coefficients satisfying (3.7). (In connection with requirement (2.10) for an essential singularity, note that if $a_0 \equiv 0$, we may set $\phi_0 \equiv 1$). Theorems II and III reduce the assertions to be proved to corresponding statements about pseudoanalytic functions established in [5].

NEW YORK UNIVERSITY.

REFERENCES.

[1] L. Bers, "Partial differential equations and generalized analytic functions," *Proceedings of the National Academy of Sciences*, vol. 36 (1950), pp. 130-136. Second note, *ibid.*, vol. 37 (1951), pp. 42-47.

[2] ———, *Theory of pseudo-analytic functions*, Lecture notes (mimeographed), New York University, 1953.

[3] ———, "Univalent solutions of linear elliptic systems," *Communications on Pure and Applied Mathematics*, vol. 6 (1953), pp. 513-526.

[4] ———, "Function-theoretical properties of solutions of partial differential equations of elliptic type," *Annals of Mathematics*, Study 33 (1954), pp. 69-94.

[5] ———, Local theory of pseudo-analytic functions, *Lectures on Functions of a Complex Variable*, 1955, pp. 213-244.

[6] L. Bers and L. Nirenberg, "On a representation theorem for linear elliptic systems with discontinuous coefficients and its applications," *Convegno Internazionale sulle Equazioni Derivate e Parziali*, Agosto 1954, pp. 111-140.

[7] A. P. Calderón and A. Zygmund, "On the existence of certain singular integrals," *Acta Mathematica*, vol. 88 (1952), pp. 85-139.

[8] K. O. Friedrichs,, "The identity of weak and strong extensions of differential operators," *Transactions of the American Mathematical Society*, vol. 55 (1944), pp. 132-151.

[9] P. Hartman and A. Wintner, "On the third fundamental form of a surface," *American Journal of Mathematics*, vol. 75 (1953), pp. 298-334.

[10] ——— and A. Wintner, "On the local behavior of solutions of non-parabolic partial differential equations," *ibid.*, vol. 75 (1953), pp. 449-476.

[11] ——— and A. Wintner, "Umbilical points and W-surfaces," *ibid.*, vol. 76 (1954), pp. 502-508.

[12] ——— and A. Wintner, "On the local behavior of solutions of non-parabolic partial differential equations. II. The uniqueness of the Green singularity," *ibid.*, vol. 76 (1954), pp. 351-361.

[13] H. Lewy, "On differential geometry in the large, I," *Transactions of the American Mathematical Society*, vol. 43 (1938), pp. 258-270.

[14] L. Nirenberg, "On nonlinear elliptic partial differential equations and Hölder continuity," *Communications on Pure and Applied Mathematics*, vol. 6 (1953), pp. 97-156.

[15] S. L. Sobolev, *Some applications of functional analysis*, Leningrad, 1950 (in Russian).

[16] I. N. Vekua, "Systems of partial differential equations of first order of elliptic type and boundary value problems with applications to the theory of shells," *Matematicheski Sbornik*, vol. 31 [73] (1952), pp. 217-314 (in Russian).

GENERALIZED LAPLACIANS.*

By Victor L. Shapiro.[1]

1. Introduction. The primary purpose of this paper is to answer the following question:

If $f(x)$ and its first r generalized Laplacians are known in a domain G where $f(x)$ and the first $r-1$ generalized Laplacians are continuous and the r-th generalized Laplacian is integrable, can $f(x)$ be obtained from its r-th generalized Laplacian by the expected integral representation?

The answer is in the affirmative, and we prove it by means of Rudin's result [5] on the first generalized Laplacian and by means of Fourier analysis. In so doing, we also solve a question in the uniqueness of multiple trigonometric series left open both by Cheng [3] and the present author [6] as well as obtain an integral depresentation akin to [5] for continuous functions defined on the torus.

2. Definitions and notation. We shall operate in n-dimensional Euclidean space $(n \geqq 2)$ designated by E_n and use the following notation:

$$m = (m_1, \cdots, m_n); x = (x_1, \cdots, x_n), \alpha x + \beta m = (\alpha x_1 + \beta m_1, \cdots \alpha x_n + \beta m_n), (m, x) = m_1 x_1 + \cdots + m_n x_n \text{ and } |x| = (x, x)^{\frac{1}{2}}.$$

As in [1], a multiple trigonometric series $\sum a_m e^{i(m,x)}$, where m represents a lattice point and a_m is an arbitrary complex number, will be said to be Abel summable at the point x to the value $L(x)$ if

$$\sum a_m \exp[i(m, x) - |m| z] \to L(x) \text{ as } z \to 0.$$

The series $\sum a_m e^{i(m,x)}$ will be called a real-valued series if $\bar{a}_m = a_{-m}$, and a series of class (U') if $\sum' a_m |m|^{-2} e^{i(m,x)}$, where $m \neq 0$, is the Fourier series of a continuous periodic function.

The open sphere of radius t and center x will be denoted in this paper by $D_n(x, t)$; the surface of this sphere by $C_n(x, t)$. The torus or fundamental semi-closed cube $\{x; -\pi < x_i \leqq \pi, i = 1, \cdots, n\}$ will be designated

* Received December 1, 1955.

[1] National Science Foundation Fellow.

by Ω_n. $x+\Omega_n$ will be the set $\{p; p-x \varepsilon \Omega_n\}$. Z contained in Ω_n will be said to be a closed set in Ω_n if it contains all its limit points in E_n except those which lie on the faces $x_i=-\pi$, $i=1,\cdots,n$. Also these latter limit points when considered as points on the faces $x_i=\pi$ do lie in Z. In other words Z will be said to be closed in Ω_n if it is closed in a torus sense. It is clear that if Z is such a set and also of capacity zero its closure in E_n is also of capacity zero.

Given $F(x)$ a real-valued function in $D_n(x_0,t_0)$ which is integrable on $C_n(x_0,t)$ for $0<t\leqq t_0$, we shall designate the mean value of F on this latter surface by $L(F,x_0,t)$. Thus

$$L(F,x_0,t)=\omega_n^{-1}\int_{C_n(0,1)}F(x_0+tx)\,dS_{n-1}(x)$$

where $\omega_n=2\pi^{\frac{1}{2}n}/\Gamma(\frac{1}{2}n)$ is the $(n-1)$-dimensional volume of $C_n(0,1)$ and $dS_{n-1}(x)$ is the $(n-1)$-dimensional volume element of $C_n(0,1)$.

We say that $F(x)$ has an r-th generalized Laplacian at the point x_0 equal to α_r if

$$L(F,x_0,t)=\Gamma(n/2)\sum_{j=0}^{r}t^{2j}\alpha_j/2^{2j}j!\Gamma(j+n/2)+o(t^{2r})$$

where the α_j $(j=0,\cdots,r)$ are constants. This r-th generalized Laplacian will be designated by $\Delta_rF(x_0)$, and it is clear that if $\Delta_rF(x_0)$ exists, then $\Delta_sF(x_0)$ exists for $0\leqq s\leqq r$. By [4, p. 261] if $F(x)$ is in class C^{2r} in $D_n(x_0,t)$ for some $t>0$, then $\Delta_rF(x_0)$ exists and equals $\Delta^rF(x_0)$, where Δ^r stands for the usual Laplace operator iterated r-times.

We set $\nabla(F,x_0,t)=L(F,x_0,t)-F(x_0)$ and

$$\psi^*F(x_0)=\limsup_{t\to 0}2n\nabla(F,x_0,t)/t^2 \tag{1}$$

the upper first generalized Laplacian of F at the point x_0. Replacing lim sup by lim inf in (1) we have a similar definition for the lower first generalized Laplacian $\psi_*F(x_0)$. If $\psi^*F(x_0)=\psi_*F(x_0)$ is finite, then F has a first generalized Laplacian at the point x_0.

Throughout this paper μ will always designate the value $(n-2)/2$, and $J_\mu(t)$ will stand for the Bessel function of the first kind of order μ.

The capacity of a set in E_n will refer to the logarithmic capacity if $n=2$ and n-dimensional Newtonian capacity if $n\geqq 3$. In this connection we construct the function $\Phi(x)$ defined in Ω_n as

$$\Phi(x)=2\pi\log|x|^{-1}\text{ for }n=2\text{ and }\Phi(x)=(2\pi)^n[\omega_n(n-2)]^{-1}|x|^{-(n-2)}\quad\text{for }n\geqq 3.$$

$\Phi(x)$ is then defined in the rest of E_n by periodicity. In other words, if $\eta = 2\pi(j_1, \cdots, j_n)$, where j_k $(k = 1, \cdots, n)$ are integers, then $\Phi(x+\eta) = \Phi(x)$.

The function whose Fourier series is $\sum' e^{i(m,x)} | m |^{-2}$, where $m \neq 0$, will be designated in this paper by $G(x)$. The properties of $G(x)$, which are the key to the whole paper, will be discussed in Section 4.

If $f(x)$ is integrable in a bounded domain R', by $\Phi * f$ we shall mean $\int_{R'} \Phi(x-y) f(y) dy$. By $F = \Phi * f$ will be meant that

$$\int_{R'} | F(x) - \Phi * f(x) | \, dx = 0.$$

$H_r(x)$ will be said to be harmonic of order r in R' if $H_r(x)$ is in class C^∞ and $\Delta^r H_r(x) = 0$.

If $f(x)$ is integrable on Ω_n, then $S[f]$ will designate the Fourier series of f.

$\bar{R}$ will designate the closure of R in E_n.

3. Statement of main results. We shall prove the following theorems:

THEOREM 1. *Let $f(x)$ and $\Delta_j f(x)$ $(j = 1, \cdots, r-1)$ be continuous in a bounded domain R contained in E_n and let Z be a closed and bounded set of capacity zero contained in E_n. Suppose that $\Delta_r f(x)$ exists in $R - RZ$ and is integrable on R. Then in every subdomain R' whose closure is contained in R,*

$$f = \underbrace{P * \cdots * P}_{r} * \Delta_r f + H_r,$$

where $H_r(x)$ is harmonic of order r in R' and $P(x) = -(2\pi)^{-1} \log | x |^{-1}$ if $n = 2$ and $= -[(n-2)\omega_n]^{-1} | x |^{-(n-2)}$ if $n \geqq 3$.

Remark. The condition that f and $\Delta_j f$ be continuous in the above theorem is necessary as can be seen from the following example in E_2. Set $f(x) = x_1^2$ if $x_1 \geqq 0$ and $= -x_1^2$ if $x_1 \leqq 0$. Then $\Delta_2 f(x) = 0$ for all x in E_2. But $f(x)$ is clearly not harmonic of order 2.

THEOREM 2. *Let $F(x)$ and $g(x)$ be two real-valued periodic functions of period 2π in each variable, with $F(x)$ continuous in E_n and $g(x)$ integrable on the torus Ω_n. Also let Z be a closed (in the torus sense) set of capacity zero contained in Ω_n. Suppose that*

(i) $\psi^* F(x) > -\infty$, $\psi_* F(x) < +\infty$ *for x in $\Omega_n - Z$;*

(ii) $g(x) \leqq \psi^* F(x)$ *in* Ω_n.

Then

a) $\Delta_1 F(x)$ *exists at almost all points* x *of* Ω_n *and is in* L_1 *on* Ω_n;

b) *at all points* x *for which* $\int_{\Omega_n} |\Delta_1 F(y)| \, |\Phi(x-y)| \, dy < \infty$, *we have*

$$F(x) = -(2\pi)^{-n} \int_{\Omega_n} \Delta_1 F(y) G(x-y) \, dy + (2\pi)^{-n} \int_{\Omega_n} F(y) \, dy.$$

THEOREM 3. *Given the multiple trigonometric series* $\sum a_m e^{i(m,x)}$ *where the* a_m *are arbitrary complex numbers. Let* Z *be a closed (in the torus sense) set of capacity zero contained in* Ω_n. *Suppose that*

(i) *the series is of class* (U');

(ii) *the series is Abel summable to* $f(x)$ *almost everywhere where* $f(x)$ *is in* L_1 *on* Ω_n;

(iii) $\limsup_{z \to 0} | \sum a_m e^{i(m,x)-|m|z} | < \infty$ *in* $\Omega_n - Z$.

Then the series is the Fourier series of $f(x)$.

4. Green's function on the torus. Before proving Theorem 1, it is necessary to investigate the properties of the function $G(x)$ whose Fourier series is $\sum' e^{i(m,x)} |m|^{-2}$. We first obtain the following lemma with $\Phi(x)$ as in Section 2.

LEMMA 1. $G(x) = \Phi(x) + H^*(x)$ *where* $H^*(x)$ *is continuous in* E_n.

Let us set λ_m^{-1} equal to the m-th Fourier coefficient of $\Phi(x)$. Then we prove the lemma by showing that $\sum_{m \neq 0} | \lambda_m^{-1} - |m|^{-2} | < \infty$.

For $n = 2$, this already has been shown in [7]. For $n \geqq 3$, we observe by Green's second identity that for $\mu = (n-2)/2$ and $m \neq 0$ that

$$(2) \qquad [\omega_n(n-2)]^{-1} \int_{\Omega_n - D_n(0,\epsilon)} |x|^{-(n-2)} e^{i(m,x)} dx$$
$$= 2^\mu \Gamma(\mu+1) J_\mu(|m|\epsilon)(|m|\epsilon)^{-\mu} |m|^{-2} + o(1)$$
$$+ K |m|^{-2} \sum_{j=1}^{n} \int_{\Omega_{n-1}} \frac{\exp[i(m_1 x_1 + \cdots + m_j x_j^* + \cdots + m_n x_n)]}{(x_1^2 + \cdots + \pi^2 + \cdots + x_n^2)^{n/2}}$$
$$\times \cos m_j \pi \, dx_1 \cdots dx_j^* \cdots dx_n$$

where K is a constant independent of m and ϵ and $m_j x_j^*$ stands for the deletion of $m_j x_j$.

We shall now show that (9) holds. For by [1, p. 176]

$$F(z,0)/z = K_n \int_0^\infty L(F,0,t)(z^2+t^2)^{-(n+1)/2} t^{n-1}\,dt$$

where K_n is a positive constant.

But since $F(x)$ is a continuous periodic function, it is not difficult to see that for every $z > 0$, there is an open interval containing z such that

$$\lim_{R\to\infty} \int_0^R zL(F,0,t)(z^2+t^2)^{-(n+3)/2} t^{n-1}\,dt$$

is uniformly convergent in this interval. Therefore

$$(10) \quad \partial[-F(z,0)/z]/\partial z = (n+1)K_n \int_0^\infty zL(F,0,t)(z^2+t^2)^{-(n+3)/2} t^{n-1}\,dt.$$

From the fact that $\psi_* F(0) = h > 0$, we have that there is a $\delta > 0$ such that for $0 < t \leqq \delta$, $L(F,0,t) > ht^2/4n$. The fact in conjunction with the observation that $\lim_{z\to 0} z \int_\delta^\infty t^{n-1}[z^2+t^2]^{-(n+3)/2}\,dt = 0$, tells us from (10) that

$$(11) \quad \liminf_{z\to 0} \partial[-F(z,0)/z]/\partial z > hK_n' \liminf_{z\to 0} \int_0^\delta z(z^2+t^2)^{-(n+3)/2} t^{n+1}\,dt,$$

where K_n' is a positive constant. But

$$\lim_{z\to 0} \int_0^\delta z(z^2+t^2)^{-(n+3)/2} t^{n+1}\,dt = \int_0^\infty t^{n+1}(1+t^2)^{-(n+3)/2}\,dt > 0,$$

and we have shown that (9) holds.

By $\Delta^r S[f]$ we shall mean the trigonometric series that one obtains by formally applying the Laplace operator r-times to $S[f]$.

We now state a lemma concerning such series whose proof follows immediately from [8, p. 225] for the two-dimensional case and from a similar theorem for the n-dimensional case.

LEMMA 5. *If $\Delta_r f(x_0)$ exists, $\Delta^r S[f]$ is at the point x_0 Abel summable to $\Delta_r f(x_0)$.*

6. Proof of Theorem 1. Since R is assumed to be a bounded domain, we shall assume further that its closure is contained in the interior of Ω_n. From the proof it will be apparent that this additional assumption will cause no loss of generality. Also we shall assume from the start that $f(x)$ is real-valued with no loss of generality.

Since $\bar{R}' \subset R$, we can insert three other domains between them with the same property as R', i.e. $\bar{R}' \subset R'' \subset \bar{R}'' \subset R''' \subset \bar{R}''' \subset R^{\text{iv}} \subset \bar{R}^{\text{iv}} \subset R$ and we can form the localizing function $\lambda(x)$ for R''' and R^{iv}. Then $\lambda(x) = 1$ for x in R''' and $= 0$ for x in Ω_n and not in R^{iv}, and furthermore $\lambda(x)$ is in class C^∞ for x in Ω_n.

By [5, Theorem 1], the theorem is true for $r = 1$. Let us assume then that the theorem is true for $1 \leqq s \leqq r-1$ and let us set $g(x) = \lambda(x)\Delta_{r-1}f(x)$, $S[g] = \sum_m b_m e^{i(m,x)}$, and $S[F] = \sum_{m \neq 0} (-1)^{r-1} b_m e^{i(m,x)} / | m |^{2(r-1)}$ where $F(x)$ is taken to be continuous. Then by Lemma 3, $\Delta_{r-1}(F - b_0 \,| x |^{2(r-1)}/k_{r-1}) = g(x)$ where $\Delta^j \,| x |^{2j} = k_j$, and consequently in R''', $\Delta_{r-1}(F - f - b_0 \,| x |^{2(r-1)}/k_{r-1} = 0$. Therefore by the inductive assumption

$$F(x) = f(x) + H_r(x) \text{ for } x \text{ in } R'' \tag{12}$$

where $H_r(x)$ is harmonic of order r in R''.

From (12) we see that in $R'' - R''Z$, $\Delta_r F(x) = \Delta_r f(x)$. Consequently by Lemma 5, $-\sum_m b_m \,| m |^2 e^{i(m,x)}$ is Abel summable to $\Delta_r f(x)$ in $R'' - R''Z$. But then by Lemma 4, $\psi^* g \geqq \Delta_r f \geqq \psi_* g$ in $R'' - R''Z$. Since g is a continuous function in R'', the conditions of [5, Theorem 1] are satisfied. So in R'

$$g = P * \Delta_r f + H_1(x) \tag{13}$$

where $H_1(x)$ is harmonic in R'.

Since $\Delta_{r-1}(F - b_0 \,| x |^{2(r-1)}/k_{r-1}) = g$, we have by the inductive assumptions that in R'

$$F = \underbrace{P * \cdots * P}_{r-1} * g + H_{r-1} + b_0 \,| x |^{2(r-1)}/k_{r-1}. \tag{14}$$

From (12), (13), and (14), we obtain that, in R',

$$f = \underbrace{P * \cdots * P}_{r} * \Delta_r f + H_R'$$

where $H_R'(x)$ is harmonic of order r, and the proof to the theorem is complete.

7. Proof of Theorem 2. By [5, Theorem 1 and 1.3.2] for almost all x in any bounded domain R contained in the plane

$$F(x) = -(2\pi)^{-1} \int_R \Delta_1 F(y) \log | x - y |^{-1} dy + h_1(x) \tag{15}$$

where $h_1(x)$ is harmonic in R. In a similar manner it can be shown that for any bounded domain R contained in E_n $(n \geqq 3)$ and for almost all x in R

$$F(x) = -[\omega_n(n-2)]^{-1} \int_R \Delta_1 F(y) \,|\, x-y \,|^{-(n-2)} dy + h_1(x) \tag{16}$$

where $h_1(x)$ is harmonic in R. In both cases $\Delta_1 F(y)$ is in L_1 on R.

We conclude, therefore, that for any x_0 in E_n there is a continuous function $F_{x_0}(x)$ which for almost all x in $D_n(x_0, \pi/4)$ is such that

$$F_{x_0}(x) = -(2\pi)^{-n} \int_{D_n(x_0, \pi/4)} \Phi(x-y) \Delta_1 F(y) \, dy \tag{17}$$

and. furthermore, for all x in $D_n(x_0, \pi/4)$

$$\Delta[F(x) - F_{x_0}(x)] = 0. \tag{18}$$

We also observe that $\int_{\Omega_n + x_0 - D_n(x_0, \pi/4)} \Phi(x-y) \Delta_1 F(y) \, dy$ is continuous in $D_n(x_0, \pi/4)$ and that for almost all x in E_n,

$$\int_{\Omega_n + x_0} \Phi(x-y) \Delta_1 F(y) \, dy = \int_{\Omega_n} \Phi(x-y) \Delta_1 F(y) \, dy.$$

These facts in conjunction with (17) and (18) tell us that there exists a continuous function $F_1(x)$ which is also periodic of period 2π in each variable such that for almost all x in E_n,

$$F_1(x) = -(2\pi)^{-n} \int_{\Omega_n} G(x-y) \Delta_1 F(y) \, dy. \tag{19}$$

But then from (18), Lemma 2, and the fact that $\Delta G(x) = 1$ for $x \neq 0$ mod Ω_n, we obtain that, for all x in $D_n(x_0, \pi/4)$,

$$\Delta[F - F_1] = (2\pi)^{-n} \int_{\Omega_n} \Delta_1 F(y) \, dy.$$

Since x_0 is arbitrary and $F - F_1$ is a periodic continuous function, it must be that $F(x) = F_1(x) +$ constant on Ω_n. That the constant agrees with that of the theorem follows from the fact that the integral of $G(x)$ over Ω_n is zero.

Part b) of the theorem follows from the fact that by [5, Theorem 1] equality holds in (17) at all points at which b) is satisfied. But then equality also holds at those points in (19), and the theorem is proved.

8. Proof of Theorem 3. There is no loss of generality if from the start we assume that the given trigonometric series is a real-valued series. For suppose the theorem is proved under this assumption. Then it is clear that both of the series $\sum (a_m + \bar{a}_{-m}) e^{i(m,x)}$ and $\sum i(a_m - \bar{a}_{-m}) e^{i(m,x)}$ are of class (U'), are Abel summable to $f(x) + \bar{f}(x)$ and $i[f(x) - \bar{f}(x)]$ almost everywhere respectively, are real-valued trigonometric series, and satisfy (iii). We would then have

$$2a_m = (a_m + \bar{a}_{-m}) + (a_m - \bar{a}_{-m})$$

$$= (2\pi)^{-n} \int_{\Omega_n} e^{i(m,x)}[f(x) + \bar{f}(x) + f(x) - \bar{f}(x)]dx$$

and the theorem would be proved in the more general case.

We also see there is no loss in generality in asuming that $a_0 = 0$. For if the theorem is proved with constant term equal to zero, we would then have $(2\pi)^{-n} \int_{\Omega_n} [f(x) - a_0]dx = 0$, and consequently the theorem is true in the more general case.

With the given series a real-valued series, with $a_0 = 0$, and with $F(x)$ the continuous periodic function whose Fourier series is $-\sum_{m \neq 0} a_m e^{i(m,x)} | m |^{-2}$, we then have by Lemma 4 and assumption (ii) of this theorem that $\psi^* F(x) \geqq f(x)$ almost everywhere. Setting $g(x) = \min[\psi^* F(x), f(x)]$, we have that $g(x)$ is integrable on Ω_n. But then by Theorem 2, $\Delta_1 F(x)$ exists almost everywhere in Ω_n, and furthermore by Lemma 5 and (ii) of the present theorem, $\Delta_1 F(x) = f(x)$ almost everywhere. We conclude, consequently from Theorem 2, that for almost all x in Ω_n,

$$F(x) = -(2\pi)^{-n} \int_{\Omega_n} f(y) G(x-y) dy. \tag{20}$$

Observing that $\int_{\Omega_n} |f(y)| dy \int_{\Omega_n} |G(x-y)| dx < \infty$, we have that for $m \neq 0$, by Fubini's theorem and (20),

$$(2\pi)^n a_m | m |^{-2} = \int_{\Omega_n} e^{-i(m,x)} F(x) dx$$

$$= \int_{\Omega_n} e^{-i(m,y)} f(y) dy [(2\pi)^{-n} \int_{\Omega_n} e^{-i(m,x-y)} G(x-y) dx]$$

$$= \int_{\Omega_n} e^{-i(m,y)} f(y) | m |^{-2} dy.$$

This last fact in conjunction with the observation that by [1] the Fourier series of $f(x)$ is Abel summable to $f(x)$ almost everywhere tells us that $\int_{\Omega_n} f(x)\,dx = 0$, which concludes the proof to the theorem.

INSTITUTE FOR ADVANCED STUDY AND RUTGERS UNIVERSITY.

REFERENCES.

[1] S. Bochner, "Summation of multiple Fourier series by spherical means," *Transactions of the American Mathematical Society*, vol. 40 (1936), pp. 175-207.

[2] ———, "Zeta functions and Green's functions for linear partial differential operators of elliptic type with constant coefficients," *Annals of Mathematics*, vol. 57 (1953), pp. 32-56.

[3] M. T. Cheng, "Uniqueness of multiple trigonometric series," *Annals of Mathematics*, vol. 52 (1950), pp. 403-416.

[4] R. Courant and D. Hilbert, *Methoden der mathematischen Physik II*, Berlin, 1937.

[5] W. Rudin, "Integral representation of continuous functions," *Transactions of the American Mathematical Society*, vol. 68 (1950), pp. 278-286.

[6] V. L. Shapiro, "An extension of results in the uniqueness theory of double trigonometric series," *Duke Mathematical Journal*, vol. 20 (1953), pp. 359-365.

[7] ———, "Logarithmic capacity of sets and double trigonometric series," *Canadian Journal of Mathematics*, vol. 6 (1954), pp. 582-592.

[8] ———, "Circular summability C of double trigonometric series," *Transactions of the American Mathematical Society*, vol. 76 (1954), pp. 223-233.

[9] A. Zygmund, *Trigonometric series*, Warsaw, 1935.

2. In the introduction, we formulated, in addition to the problem (P), a more precise problem (P′). We may modify the problem (A) similarly, by requiring f to be everywhere biregular; call (A′) this modified problem. For (A′) to have a solution, it is obviously necessary that the $f_{\tau,\sigma}$ should be everywhere biregular and satisfy the conditions in Theorem 1.

Assume that this is so; let (V_0, f) be a solution of (A). Then, if (V_0', f') is a solution of (A′), the unicity of the solution of (A) shows that we must have $f' = f \circ F^{-1}$, where F is a birational correspondence between V_0 and V_0', defined over k_0. Thus problem (A′) may be reformulated as follows:

(B) *Let k and k_0 be as in problem* (A); *let V and V_0 be varieties, respectively defined over k and over k_0; let f be a birational correspondence, defined over k, between V_0 and V. Find a variety V_0' and a birational correspondence F between V_0 and V_0', both defined over k_0, such that the birational correspondence $f \circ F^{-1}$ between V_0' and V is everywhere biregular.*

It is obvious that, if (B) has a solution, this is unique up to a k_0-isomorphism; therefore the same is true for (A′). If (B) has a solution, then ($\mathfrak{I}$ being defined as before) the birational correspondence $f^\tau \circ (f^\sigma)^{-1}$ between V^σ and V^τ must be everywhere biregular for all σ, τ in $\mathfrak{I}$. We will prove that this condition is also sufficient, at any rate if V is a k-open subset of a projective variety, defined over k. This will be an immediate consequence of the following result.

PROPOSITION 1. *Let k, k_0, $\mathfrak{I}$ be as in* (A). *Let V_0 be a variety, defined over k_0; let V be a projective* (*resp. affine*) *variety, defined over k; let f be a birational correspondence, defined over k, between V_0 and V. Then there is a projective* (*resp. affine*) *variety W and a birational correspondence F between V_0 and W, both defined over k_0, such that $F \circ f^{-1}$ is biregular at every point of V where the mappings $f^\sigma \circ f^{-1}$ are defined for all $\sigma \in \mathfrak{I}$.*

Let S be the ambient space of V, projective or affine; f may be regarded as a mapping of V_0 into S. Call $\sigma_1 = \epsilon, \sigma_2, \cdots, \sigma_n$ the elements of $\mathfrak{I}$, and put $F_1 = (f^{\sigma_1}, \cdots, f^{\sigma_n})$; this is a mapping of V_0 into the product $S \times \cdots \times S$ of n factors equal to S, and is defined over the compositum K of the fields k^σ. It is clear that $F_1 \circ f^{-1}$ is defined wherever all the $f^\sigma \circ f^{-1}$ are defined. Let x be a generic point of V_0 over k_0; let W_1 be the locus of $F_1(x)$ over K; put $u = F_1(x) = (x_1, \cdots, x_n)$. As $\sigma_1 = \epsilon$, we have $x_1 = f(x)$, so that the image of u by the mapping $f \circ F_1^{-1}$ is x_1; this shows that $f \circ F_1^{-1}$ is the mapping induced on W_1 by the projection of the product $S \times \cdots \times S$ onto its first factor, and is therefore everywhere defined. Thus the birational correspon-

dence $F_1 \circ f^{-1}$ between V and W_1 is biregular wherever all the $f^\sigma \circ f^{-1}$ are defined.

Let $z_1, \cdots, z_n$ be n points of S; if S is the projective m-space, put $z_i = (z_{i0}, \cdots, z_{im})$; and let z' be the point, in a projective space of suitable dimension, whose homogeneous coordinates are all the monomials $z_{1\mu_1} z_{2\mu_2} \cdots z_{n\mu_n}$, with $0 \leqq \mu_i \leqq m$ for every i. If S is the affine m-space, put $z_i = (z_{i1}, \cdots, z_{im})$, put $z_{i0} = 1$ for $1 \leqq i \leqq n$, and let z' be the point, in an affine space of suitable dimension, whose coordinates are the same monomials as before. In either case, put $z' = \Phi(z_1, \cdots, z_n)$; it is well-known that Φ is an everywhere biregular mapping of $S \times \cdots \times S$ onto its image in projective (resp. affine) space. Put now $F_2 = \Phi \circ F_1$; then F_2 is a birational correspondence between V_0 and $W_2 = \Phi(W_1)$, and $F_2 \circ f^{-1}$ is biregular wherever all the $f^\sigma \circ f^{-1}$ are defined.

If S is projective, let $(1, f_1(x), \cdots, f_m(x))$ be a set of homogeneous coordinates for $f(x)$; the f_μ are functions on V_0, defined over k. Put $f_0 = 1$. Then we have $F_2 = (g_0, \cdots, g_r)$, where the g_ρ are all the monomials

$$f_{\mu_1}{}^{\sigma_1} f_{\mu_2}{}^{\sigma_2} \cdots f_{\mu_n}{}^{\sigma_n}.$$

If ω is an automorphism of K over k, $g_\rho{}^\omega$ is again one of the g_ρ, which we may write as $g_{\omega(\rho)}$; the mapping $\rho \to \omega(\rho)$ determines a representation of Γ (the Galois group of K over k) as a group of permutations on the g_ρ. For a given ρ, let γ_ρ be the subgroup of Γ determined by $\omega(\rho) = \rho$; then, for $\omega \varepsilon \Gamma$, $\omega(\rho)$ takes a number of distinct values equal to the index d_ρ of γ_ρ in Γ. If K_ρ is the subfield of K consisting of the elements of K invariant under γ_ρ, g_ρ is defined over K_ρ; therefore, if $(\alpha_1, \cdots, \alpha_{d_\rho})$ is a basis of K_ρ over k_0, we may write $g_\rho = \sum_\nu \alpha_\nu h_{\rho\nu}$, where the $h_{\rho\nu}$, for $1 \leqq \nu \leqq d_\rho$, are functions on V_0, defined over k_0. Then we have, for all $\omega \varepsilon \Gamma$:

$$g_{\omega(\rho)} = \sum_\nu \alpha_\nu{}^\omega h_{\rho\nu}.$$

If, in this relation, we take for ω a set of representatives of the d_ρ cosets of γ_ρ in Γ, we get a linear substitution expressing the d_ρ distinct functions $g_{\omega(\rho)}$ in terms of the d_ρ functions $h_{\rho\nu}$; and, since K_ρ is separable over k_0, that substitution is invertible. From this it follows immediately that, if we call $F(x)$ the point whose homogeneous coordinates are all the functions $h_{\rho\nu}$ (where ρ runs through a set of representatives for the classes of equivalence determined by the permutation group Γ on the set $\{0, 1, \cdots, r\}$, and where, for each ρ, we take $1 \leqq \nu \leqq d_\rho$), F is of the form $\Psi \circ F_2$, where Ψ is an automorphism of the ambient projective space of W_2. If S is affine, we put

$f = (f_1, \cdots, f_m)$, $f_0 = 1$, and we define F by the same formulas as in the projective case, but regard it as a mapping of V_0 into an affine space; then we have again $F = \Psi \circ F_2$, Ψ being now an automorphism of the ambient affine space of W_2. In either case, the mapping F is defined over k_0; if W is the locus of $F(x)$ over k_0, W and F have the properties required by our proposition.

3. Before applying this to the problems (A′) and (B), we need a general lemma:

LEMMA 1. *Let f be a birational correspondence between two varieties U and V; let k be a field of definition for U, V, f. Then the sets of points where f and f^{-1} are respectively biregular are k-open, and f determines a k-isomorphism between them.*

Call U' the set of points of U where f is defined, U'' the set of points of U where it is biregular; call V' the set of points of V where f^{-1} is defined, V'' the set where it is biregular. By [**4**], App., Prop. 8, U' and V' are k-open. Call f' the restriction of f to $U' \times V'$ (i.e. the birational correspondence between U' and V' whose graph is the set-theoretic intersection of the graph of f with $U' \times V'$). If f is biregular at a point a of U, it is defined at a, so that $a \varepsilon U'$; and, if we put $b = f(a)$, f^{-1} is defined at b, so that $b \varepsilon V'$; therefore (a, b) is on the graph of f', and f' is defined at a. Conversely, let a be a point of U' where f' is defined; put $b = f'(a)$; then b is in V', so that f^{-1} is defined at b; as U' and V' are open, f is then defined at a, and we have $b = f(a)$; thus f is biregular at a. This shows that U'' is the set of points of U' where f' is defined; similarly V'' is the set of points of V' where f'^{-1} is defined; this implies that they are k-open. If f'' is the restriction of f to $U'' \times V''$, it is everywhere biregular by definition (i.e., f'' is biregular at every point of U'', and f''^{-1} is so at every point of V'').

In order to formulate our results on problems (A′) and (B), we will say that a variety U, defined over a field k, is *projectively* (resp. *affinely*) *embeddable over* k if it is k-isomorphic to a k-open subset of a projective (resp. affine) variety, defined over k.

THEOREM 2. *Problem* (B) *has a solution* (V_0', F') *provided V is projectively embeddable over k and the birational correspondence $f^\sigma \circ f^{-1}$ between V and V^σ is everywhere biregular for every isomorphism σ of k over k_0 into $\bar{k}_0$. When that is so, V_0' is projectively embeddable over k_0; it is affinely embeddable over k_0 if V is so over k.*

We may assume V to be a k-open subset of a projective (resp. affine) variety, defined over k. Take W and F as in Proposition 1; then $F \circ f^{-1}$ is biregular at every point of V; therefore, by Lemma 1, it is a k-isomorphism between V and the k-open subset V_0' of W where $f \circ F^{-1}$ is biregular. As the $f^\sigma \circ f^{-1}$ are everywhere biregular V_0' is also the subset of W where $f^\sigma \circ F^{-1}$ is biregular, for every σ; therefore it is invariant under all automorphisms of $\bar{k}_0$ over k_0, so that it is k_0-open, by [**4**], App., Prop. 9. Then, if F' is the restriction of F to $V_0 \times V_0'$, (V_0', F') is a solution of (B).

THEOREM 3. *Problem* (A′) *has a solution, i.e., problem* (A) *has a solution* (V_0, f) *for which* f *is everywhere biregular, provided* V *is projectively embeddable over* k *and the* $f_{\tau,\sigma}$ *are everywhere biregular and satisfy the conditions in Theorem* 1. *When that is so,* V_0 *is projectively embeddable over* k_0; *it is affinely embeddable over* k_0 *if* V *is so over* k. *The solution is unique up to a* k_0*-isomorphism.*

Section II. Regular Extensions of the Groundfield.

4. Let now k denote the groundfield. Let T be a variety, defined over k; let t be a generic point of T over k. When we denote by V_t a variety, defined over $k(t)$, we will agree, whenever t' is also a generic point of T over k, to denote by $V_{t'}$ the transform of V_t by the isomorphism of $k(t)$ onto $k(t')$ over k which maps t onto t'. Similarly, if a mapping, defined over $k(t)$, is denoted by f_t, $f_{t'}$ will denote its transform by the same isomorphism; if t, t', t'' are three independent generic points of T over k, and $f_{t',t}$ is a mapping, defined over $k(t, t')$, we denote by $f_{t'',t'}$ the transform of $f_{t',t}$ by the isomorphism of $k(t, t')$ onto $k(t', t'')$ over k which maps (t, t') onto (t', t''); etc.

Let V_t be a variety, defined over $k(t)$; assume that there is a variety V, defined over k, and a birational correspondence f_t, defined over $k(t)$, between V and V_t; then $f_{t'} \circ f_t^{-1}$ is a birational correspondence between V_t and $V_{t'}$. We therefore modify problem (P) of the introduction as follows:

(C) *Let* T *be a variety, defined over a field* k; *let* t, t' *be independent generic points of* T *over* k. *Let* V_t *be a variety, defined over* $k(t)$; *let* $f_{t',t}$ *be a birational correspondence, defined over* $k(t, t')$, *between* V_t *and* $V_{t'}$. *Find a variety* V, *defined over* k, *and a birational correspondence* f_t, *defined over* $k(t)$, *between* V *and* V_t, *such that* $f_{t',t} = f_{t'} \circ f_t^{-1}$.

THEOREM 4. *Problem* (C) *has a solution if and only if* $f_{t',t}$ *satisfies the condition:*

(i) $$f_{t'',t} = f_{t'',t'} \circ f_{t',t}$$

where t'' is a generic point of T over $k(t,t')$. When that is so, the solution is unique, up to a birational transformation on V, defined over k.

The condition is obviously necessary. The proof for the unicity of the solution, when one exists, is quite similar to the proof of the corresponding statement in Theorem 1. Now, assuming (i) to be fulfilled, we shall construct a solution of (C). We may replace T by any birational transform of T over k, and so we may assume that T is an affine variety. Similarly we may assume that V_t is an affine variety; and, taking x to be a generic point of V_t over $k(t)$, we may replace x by (x,t) and V_t by the locus of (x,t) over $k(t)$; after that is done, V_t is still an affine variety, and we have $k(t) \subset k(x)$; from now on, assume that this is so, and assume that x has been taken generic on V_t over $k(t,t',t'')$. By [**4**], App., Prop. 1, $k(x)$ is a regular extension of k; call X the locus of x over k. Put $x' = f_{t',t}(x)$; this is a generic point of $V_{t'}$ over $k(t,t',t'')$; by the definition of $V_{t'}$, this implies that there is an isomorphism of $k(t,x)$ onto $k(t',x')$ over k, mapping t onto t' and x onto x'; therefore we have $k(t') \subset k(x')$, hence $k(x,t') \subset k(x,x')$. As the definition of x' shows $k(x,x')$ to be contained in $k(t',t,x)$, i.e. in $k(x,t')$, it follows that we have $k(x,x') = k(x,t')$; therefore x' has a locus W_x over $k(x)$. Let $k(v)$ be the smallest field of definition containing k for W_x; as $k(v) \subset k(x)$, $k(v)$ is a regular extension of k. Call V the locus of v over k; we may write $v = G(x)$, where G is a mapping of X into V, defined over k.

If we put $x'' = f_{t'',t}(x)$, W_x is also the locus of x'' over $k(x)$; as the fields $k(x,x')$ and $k(x,x'')$ are respectively the same as $k(x,t')$ and $k(x,t'')$ and are therefore algebraically independent over $k(x)$, W_x is also the locus of x'' over $k(x,x')$. But (i) may be written $x'' = f_{t'',t'}(x')$; therefore $W_{x'}$ is the same as W_x. This implies that the isomorphism of $k(x)$ onto $k(x')$ over k which maps x onto x' leaves invariant all the elements of the smallest field of definition of W_x, hence also all the elements of $k(v)$, so that we have $G(x) = G(x')$.

On the other hand, let K be an overfield of k, algebraically independent from $k(x,x')$ over k; if ϕ is any function on X, defined over K, it will induce on W_x a function which is defined over $K(v)$; if $\phi(x) = \phi(x')$, that function is a constant, so that its constant value must be in $K(v)$. This shows that $K(v)$ is the subfield of $K(x)$ consisting of the elements of $K(x)$ which are invariant under the isomorphism of $K(x)$ onto $K(x')$ over K mapping x onto x'.

Now the relation $x'' = f_{t'',t}(x)$ shows that x'' is rational over $k(t'',t,x)$,

i. e. over $k(t'', x)$, so that we may write $x'' = \phi_{t''}(x)$, where $\phi_{t''}$ is a mapping of X into $V_{t''}$, defined over $k(t'')$. The relation $x'' = f_{t'',t'}(x')$ may then be written as $x'' = \phi_{t''}(x')$. Applying to the field $K = k(t'')$ and to the function $\phi_{t''}$ what we have proved above, we conclude from this that $k(x'') \subset k(t'', v)$. As we have $G(x) = G(x')$, hence also $G(x) = G(x'')$, the isomorphism of $k(x)$ onto $k(x'')$ over k which maps x onto x'' leaves $v = G(x)$ invariant; applying the inverse of that isomorphism to the relation $k(x'') \subset k(t'', v)$, we get $k(x) \subset k(t, v)$, hence $k(x) = k(t, v)$ since $k(t)$ and $k(v)$ are both contained in $k(x)$. Also, since $k(x)$ and $k(t')$ are algebraically independent over k, the same is true of $k(v)$ and $k(t')$; as the isomorphism of $k(x')$ onto $k(x)$ over k which maps x' onto x maps t' onto t and v onto itself, this implies that $k(v)$ and $k(t)$ are algebraically independent over k. As the relation $k(x) = k(t, v)$ can also be written $k(t, x) = k(t, v)$, we conclude that V_t and V are birationally equivalent over $k(t)$, so that we may write $x = f_t(v)$, where f_t is a birational correspondence between V and V_t, defined over $k(t)$. Then we have $x' = f_{t'}(v)$. Therefore (V, f_t) is a solution of our problem. We also see that X is birationally equivalent to $T \times V$ over k.

5. Just as in Section I, we consider the problem (C′) which consists in finding a solution (V, f_t) of (C) such that f_t is everywhere biregular. For such a solution to exist, it is necessary that $f_{t',t}$ should be everywhere biregular; it will be shown that this is sufficient.

As in Section I, if we make use of Theorem 4, we see that (C′) may be reformulated as follows:

(D) *Let k, T and t be as in* (C); *let V and V_t be varieties, respectively defined over k and over $k(t)$; let f_t be a birational correspondence, defined over $k(t)$, between V and V_t. Find a variety V' and a birational correspondence F between V and V', both defined over k, such that the birational correspondence $f_t \circ F^{-1}$ between V' and V_t is everywhere biregular.*

In order to solve (D), we need some preliminary results.

LEMMA 2. *Let F and H be mappings of a variety X into two varieties W, T, all these being defined over a field k; x being a generic point of X over k, assume that $t = H(x)$ is generic over k on T and that x has a locus V_t over $k(t)$. Let F_t be the mapping of V_t into W induced by F on V_t. Then F is defined at every point of V_t where F_t and H are both defined.*

It is clearly enough to treat the case in which X is an affine variety and W is the affine line. Then F_t is the function on V_t, defined over $k(t)$,

such that $F_t(x) = F(x)$. If F_t is defined at a point a of V_t, we can write it as $F_t(x) = P(x)/Q(x)$, where P, Q are polynomials with coefficients in $k(t)$, such that $Q(a) \neq 0$. More explicitly, we have

$$P(X) = \sum_i \lambda_i(t) M_i(X), \qquad Q(X) = \sum_j \mu_j(t) N_j(X),$$

where the λ_i, μ_j are functions on T, defined over k, and the M_i, N_j are monomials in the indeterminates (X); and we have

$$(1) \qquad \sum_j \mu_j(t) N_j(a) \neq 0.$$

Then we have $F(x) = \Phi(x)/\Psi(x)$, with

$$(2) \qquad \Phi(x) = \sum_i \lambda_i(H(x)) M_i(x), \qquad \Psi(x) = \sum_j \mu_j(H(x)) N_j(x).$$

As a is on V_t, (t, a) is a specialization of (t, x) over k; if H is defined at a, we must have $H(a) = t$. As t is generic on T over k, the functions λ_i, μ_j are defined at t; therefore the functions $\lambda_i \circ H$, $\mu_j \circ H$ are defined at a on X, with the values $\lambda_i(t)$, $\mu_j(t)$. That being so, the relations (1), (2) show that F is defined at a on V.

PROPOSITION 2. *Let k, T, t, t' be as in* (C); *let V be a variety, defined over k; let V_t be a variety, defined and projectively* (*resp. affinely*) *embeddable over $k(t)$; let f_t be a birational correspondence, defined over $k(t)$, between V and V_t. Then:*

(i) *if a is a point of V_t where $f_{t'} \circ f_t^{-1}$ is biregular, there is an affine variety W and a birational correspondence F between V and W, both defined over k, such that $F \circ f_t^{-1}$ is biregular at a;*

(ii) *if $f_{t'} \circ f_t^{-1}$ is everywhere biregular, there is a variety W, defined and projectively* (*resp. affinely*) *embeddable over k, and a birational correspondence F between V and W, defined over k, such that $F \circ f_t^{-1}$ is everywhere biregular.*

We may assume that V_t is a $k(t)$-open subset of a variety, defined over $k(t)$, in a projective (resp. affine) space S. We may also assume that T is a projective (resp. affine) variety; let S' be its ambient space. If S, S' are affine, $S \times S'$ is an affine space; if they are projective, call Φ the well-known biregular embedding of $S \times S'$ into a projective space S'' of suitable dimension. Let v be generic on V over $k(t, t')$, and put $x = f_t(v)$. We may replace V_t by a suitable $k(t)$-open subset of the locus of (x, t) over $k(t)$ in the affine case, of $\Phi(x, t)$ over $k(t)$ in the projective case; after that is done,

we have $k(t) \subset k(x)$, and therefore $k(x) = k(t,x) = k(t,v)$, so that x has a locus X over k, birationally equivalent to $T \times V$, and that we may write $t = H(x)$, where H is a mapping of X into T, defined over k; moreover, the mapping H is everywhere defined on X.

Now, since X is birationally equivalent to $T \times V$ over k, and V is birationally equivalent to V_t over $k(t)$, X is birationally equivalent to $T \times V_t$ over $k(t)$. More explicitly, if we put $x' = f_{t'}(v)$, x' is generic over $k(t)$ on X, and we have $k(x') = k(t',v)$, hence $k(t,x') = k(t,t',x)$, so that we may write $x' = g_t(t',x)$, where g_t is a birational correspondence, defined over $k(t)$, between $T \times V_t$ and X. We have $t' = H(x')$, and we may write $x = \phi_t(x')$; where ϕ_t is a mapping of X into V_t, defined over $k(t)$; then (H, ϕ_t) is the mapping of X into $T \times V_t$, inverse to g_t. The mapping g_t induces on the subvariety $t' \times V_t$ of $T \times V_t$ the mapping $(t',x) \rightarrow x' = f_{t'}(f_t^{-1}(x))$; and ϕ_t induces on $V_{t'}$ the mapping $x' \rightarrow x$, i.e. the mapping $f_t \circ f_{t'}^{-1}$. Applying Lemma 2, we see that g_t is defined at (t',a) whenever a is a point of V_t where $f_{t'} \circ f_t^{-1}$ is defined, and that ϕ_t is defined at every point of $V_{t'}$ where $f_t \circ f_{t'}^{-1}$ is defined. Therefore g_t is biregular at (t',a) whenever a is a point of V_t where $f_{t'} \circ f_t^{-1}$ is biregular.

Now let A_0 be the $k(t)$-closed subset of $T \times V_t$ where g_t is not biregular; and assume first that a is a point of V_t with the property stated in (i). Then (t',a) is not in A_0, so that $T \times a$ is not contained in A_0; let A_1 be the (non-dense) $k(t,a)$-closed subset of T consisting of those points t_1 such that $(t_1,a) \varepsilon A_0$. By [4], App., Prop. 12, there is a $\bar{k}$-closed subset A_2 of T containing all $\bar{k}$-closed subsets of T contained in A_1; in particular, every point of A_1 which is algebraic over k must be in A_2. Let A_3 be the union of the components of A_2 and of their conjugates over k; put $T' = T - A_3$; this is a k-open subset of T such that, if t_1 is any algebraic point over k in T', g_t is biregular at (t_1,a).

On the other hand, assume, as in (ii), that $f_{t'} \circ f_t^{-1}$ is everywhere biregular. Then g_t is biregular at every point of $t' \times V_t$, so that A_0 has no point in common with $t' \times V_t$. This implies that the projection of A_0 on T is non-dense in T, so that, if we call A_1' the closure of that projection, it is a (non-dense) $k(t)$-closed subset of T. Let A_2' be the maximal $\bar{k}$-closed subset of T contained in A_1'; let A_3' be the union of the components of A_2' and of their conjugates over k; put $T'' = T - A_3'$. Then T'' is k-open on T; and, if t_1 is any algebraic point over k in T'', g_t is biregular at every point of $t_1 \times V_t$.

Now let t_1 be a separably algebraic point over k in T' (resp. T''); if k is finite, we may take for t_1 any algebraic point over k in T' (resp. T''),

since in that case every algebraic extension of k is separable; if k is infinite, we apply [4], App., Prop. 13. Let $t_1, \cdots, t_n$ be the distinct conjugates of t_1 over k. As they are in T' (resp. T''), g_t is biregular at (t_i, a) (resp. at every point of $t_i \times V_t$), and a fortiori at (t_i, x), for $1 \leqq i \leqq n$; therefore it induces on $t_i \times V_t$ a birational correspondence g_i between V_t and the locus V_i of the point $g_i(x) = g_t(t_i, x)$ over $k(t, t_i)$ in the projective (resp. affine) ambient space of X; and g_i is biregular at a (resp. at every point of V_t). But, as we have already observed, the relation $k(x) = k(t, v)$ shows that X is birationally equivalent to $T \times V$; we may write $x = f(t, v)$, where f is a birational correspondence between $T \times V$ and X, defined over k; then we have $x' = f(t', v)$; and f is the product of g_t and of the birational correspondence $(t', v) \rightarrow (t', x)$ between $T \times V$ and $T \times V_t$. As the latter correspondence is biregular at (t_i, v), and g_t is biregular at (t_i, x), for $1 \leqq i \leqq n$, we see that f is biregular at (t_i, v), and that we have

$$g_i(x) = g_t(t_i, x) = f(t_i, v).$$

As the point $f(t_i, v)$ has the same locus over $k(t_i)$ as over $k(t, t_i)$, this shows that V_i is defined over $k(t_i)$. As every automorphism of $\bar{k}$ over k can be extended to an automorphism of $\bar{k}(v)$ over $k(v)$, this also shows that V_i is the transform of V_1 by the isomorphism of $k(t_1)$ onto $k(t_i)$ over k which maps t_1 onto t_i. Also, if f_i is the mapping of V into V_i, defined over $k(t_i)$, which is such that $f_i(v) = f(t_i, v)$, we have $f_i = g_i \circ f_t$; and f_i is the transform of f_1 by the isomorphism of $k(t_1)$ onto $k(t_i)$ over k which maps t_1 onto t_i.

Now apply Proposition 1 to the variety V, defined over the groundfield k, to the variety V_1, defined over $k(t_1)$, and to the birational correspondence f_1; this gives a projective (resp. affine) variety W and a birational correspondence F between V and W, both defined over k, such that $F \circ f_1^{-1}$ is biregular wherever all the $f_i \circ f_1^{-1}$ are defined, i. e. wherever all the $g_i \circ g_1^{-1}$ are defined. Now, in case (i), all the g_i are biregular at a, so that all the $g_i \circ g_1^{-1}$ are biregular at the point $g_1(a)$; therefore $F \circ f_t^{-1}$, which is the same as $(F \circ f_1^{-1}) \circ g_1$, is biregular at a; as this involves merely a local property of W at the image of a by that mapping, we may replace W, in the projective case, by one of its affine representatives. Thus we have solved our problem in case (i). In case (ii), g_i is biregular at every point of V_t; as we have just shown, this implies that $F \circ f_t^{-1}$ is biregular at every point of V_t, so that it determines an isomorphism of V_t onto a $k(t)$-open subset W' of W. The assumption in (ii) implies that W' is invariant under the isomorphism of $k(t)$ onto $k(t')$ over k which maps t onto t'. From this and from [4], App., Prop. 9, it follows easily that W' is k-open; thus (W', F) is a solution of our problem.

COROLLARY. *Let k, T, t and t' be as in* (C); *let V be a variety, defined over k; let V_t be a variety, defined over $k(t)$; let f_t be a birational correspondence between V and V_t, defined over $k(t)$ and such that $f_{t'} \circ f_t^{-1}$ is everywhere biregular. Then, if a is any point of V_t, there is an affine variety W and a birational correspondence F between V and W, both defined over k, such that $F \circ f_t^{-1}$ is biregular at a.*

We may assume that t' has been taken generic on T over $k(t, a)$; take t'' generic on T over $k(t, t', a)$. Call a', a'' the images of a by $f_{t'} \circ f_t^{-1}$ and by $f_{t''} \circ f_t^{-1}$, respectively. The isomorphism of $k(t, a, t')$ onto $k(t, a, t'')$ over $k(t, a)$ which maps t' onto t'' maps a' onto a''; therefore, if $V_{t'\alpha}$ is a representative of the (abstract) variety $V_{t'}$ on which a' has a representative a_α', the point a'' of $V_{t''}$ has a representative a_α'' on $V_{t''\alpha}$. Let $f_{t'\alpha}$ be the birational correspondence between V and $V_{t'\alpha}$ which is determined by $f_{t'}$. As $f_{t''} \circ f_{t'}^{-1}$ is everywhere biregular and maps a' onto a'', $f_{t''\alpha} \circ f_{t'\alpha}^{-1}$ is biregular at a_α'. Applying Proposition 2(i) to V, $V_{t'\alpha}$ and $f_{t'\alpha}$, we get a solution (W, F) of our problem.

6. Now we can deal with problems (D) and (C').

THEOREM 5. *Problem* (D) *has a solution if and only if $f_{t'} \circ f_t^{-1}$ is everywhere biregular for t' generic over $k(t)$ on T.*

The condition being obviously necessary, assume that it is fulfilled. By the corollary of Proposition 2, there is, to every point a of V_t, an affine variety W_a and a birational correspondence F_a between V and W_a, both defined over k, such that $F_a \circ f_t^{-1}$ is biregular at a; call Ω_a the $k(t)$-open subset of V_t where $F_a \circ f_t^{-1}$ is biregular, and call W_a' its image on W_a by $F_a \circ f_t^{-1}$, which is a $k(t)$-open subset of W_a. Then W_a' is the subset of W_a where $f_t \circ F_a^{-1}$ is biregular; as in the proof of Proposition 2, this implies that W_a' is invariant under the isomorphism of $k(t)$ onto $k(t')$ over k which maps t onto t', and we again conclude from this that W_a' is k-open. As we have $a \in \Omega_a$ for every $a \in V_t$, the open sets Ω_a form a covering of V_t; by the well-known "compactoid" property of open sets in the Zariski topology, there must be finitely many points a_α on V such that the sets Ω_{a_α} cover V_t. Then the k-open subsets W_{a_α}' of the affine varieties W_{a_α}, together with the birational correspondences $F_{a_\beta} \circ F_{a_\alpha}^{-1}$ between them, define an abstract variety, which, together with the obvious birational correspondence between it and V, solves our problem.

THEOREM 6. *Problem* (C') *has a solution, i.e., problem* (C) *has a*

solution (V, f_t) *for which* f_t *is everywhere biregular, if and only if* $f_{t',t}$ *is everywhere biregular and satisfies condition* (i) *in Theorem* 4. *The solution is unique up to a* k*-isomorphism.*

This is an immediate consequence of Theorems 4 and 5.

7. As to the projective or affine embeddability of the solution of problems (D) and (C′), we have the following result.

THEOREM 7. *Let* V *be a variety, defined over a field* k, *and projectively (resp. affinely) embeddable over an overfield* K *of* k. *Then* V *is projectively (resp. affinely) embeddable over* k *provided* (i) K *is separable over* k *or* (ii) V *is everywhere normal with reference to* k.

The assumption means that there is a birational correspondence f, defined over K and biregular at every point of V, between V and a subvariety of a projective (resp. affine) space; if we regard f as a mapping of V into that space, it has a smallest field of definition k' containing k; we may replace K by k'; after that is done, K is finitely generated over k. If K is separable over k, it is a regular extension $k_1(t)$ of the algebraic closure k_1 of k in K, and k_1 is a separably algebraic extension of k of finite degree. Proposition 2(ii) shows that V is then projectively (resp. affinely) embeddable over k_1; by Theorem 2, this implies that the same is true over k; this completes the proof in case (i). If K is not separable over k, let k^* be the union of the fields $k^{p^{-n}}$, for $n = 1, 2, \cdots$; then the compositum K^* of K and k^* is separable over k^*, so that, by what we have just proved, V is projectively (resp. affinely) embeddable over k^*. In order to deal with case (ii), it is therefore enough to prove our theorem in the case in which V is everywhere normal with reference to k, and K is purely inseparable over k; I owe the proof for this to T. Matsusaka; it is as follows.

We may again assume that K is finitely generated over k; as it is purely inseparable, it is contained in some field $k' = k^{1/q}$, where q is a power of the characteristic. Then there is a mapping f' of V into a projective (resp. affine) space, defined over k', such that f' determines a birational correspondence, biregular at every point of V, between V and the closure W' of its image by f'; then W' is a projective (resp. affine) variety, defined over k', and f' determines a k'-isomorphism between V and a k'-open subset of W'. Call π the automorphism $\xi \rightarrow \xi^q$ of the universal domain; put $W = W'^{\pi}$; W is then a projective (resp. affine) variety, defined over k. Let x be a generic point of V over k; then W' is the locus of the point $y' = f'(x)$ over k', and W is

the locus of the point $y = y'^{\pi}$ over k. As y' is rational over $k'(x)$, y is so over $k(x^{\pi})$; we may write $y = g(x)$, where g is a mapping of V into W, defined over k; as we have $k'(y') = k'(x)$, we have $k(y) = k(x^{\pi})$, which implies that $k(x)$ is purely inseparable over $k(y)$. In the projective case, let U be the projective variety derived from W by normalization in the field $k(x)$ [1]; U is birationally equivalent to V over k; let z be the point of U which corresponds to x on V. In the affine case, we take for z a point in a suitable affine space such that $k[z]$ is the integral closure of the ring $k[y]$ in the field $k(x)$, and for U the locus of z over k. In either case we may write $z = f(x)$, where f is a birational correspondence between V and U, defined over k. By definition, U is everywhere normal with reference to k, and the mapping $h = g \circ f^{-1}$ of U into W is everywhere defined and such that the (set-theoretic) inverse image of every point of W for that mapping consists of finitely many points of U. Let a be any point of V; let (a, b) be a specialization of (x, z) over $x \rightarrow a$ with reference to k; then, as h is defined at b, $(a, b, h(b))$ is a specialization of (x, z, y) over k. As f' is defined at a, g is also defined there, so that we must have $h(b) = g(a)$; therefore b is one of the finitely many points of U whose image by h is $g(a)$. As V is normal at a by assumption, with reference to k, this implies that f is defined at a, and that we have $b = f(a)$. We have $g(a) = f'(a)^{\pi}$, hence $f'(a) = g(a)^{\pi^{-1}}$; as $g(a) = h(b)$, this shows that $f'(a)$ is the unique specialization of y' over $z \rightarrow b$ with reference to k'; as f' is biregular at a, f'^{-1} is defined at $f'(a)$, and therefore x has no other specialization than a over $z \rightarrow b$ with reference to k', hence also with reference to k by F-II_1, Prop. 3. As U is normal at b, with reference to k, this implies that f^{-1} is defined at $b = f(a)$. We have thus shown that f is biregular at every point of V, so that it is a k-isomorphism between V and a k-open subset of U.

As a special case (already contained in Proposition 2), we see that, in problem (D), V' is projectively (resp. affinely) embeddable over k if V_t is so over $k(t)$; similarly, in problem (C'), V is projectively (resp. affinely) embeddable over k if V_t is so over $k(t)$.

8. In [4], the construction carried out in Nos. 7-9 can be advantageously replaced by the application of our Theorem 6 to the situation described in No. 6 of that paper. The application is entirely straightforward, so that no further details need be given; this shows that the recourse to the Lang-Weil

[1] U is the "derived normal model of W in the field $k(x)$" according to Zariski's definition ([5], pp. 69-70); cf. also [3].

Theorem, i. e., in substance, to the so-called "Riemann hypothesis" in the case of a finite groundfield (loc. cit., p. 374) was unnecessary; so is the assumption of normality in the final result (loc. cit., p. 375); normality had to be assumed there merely because of the use made of the Chow point in the construction on p. 370, whereas in the present paper a different device was adopted (in the proof of Proposition 1). Of course, in the main theorem of [**4**] (p. 375), parts (i) and (ii) remain unchanged. For the sake of completeness, we give here the improved result by which part (iii) of that theorem may now be replaced:

PROPOSITION 3. *Let G be a group and W a chunk of transformation-space with respect to G, both defined over k. Then there is a transformation-space S with respect to G, and a birational correspondence f between W and S, both defined over k, with the following properties:* (a) *f is biregular at every point of W;* (b) *for every $s \varepsilon G$ and $a \varepsilon W$ such that sa is defined, we have $f(sa) = sf(a)$;* (c) *every point of S can be written in the form $sf(a)$, with $s \varepsilon G$ and $a \varepsilon W$. Moreover, S is uniquely determined by these properties up to a k-isomorphism compatible with the operations of G.*

UNIVERSITY OF CHICAGO.

REFERENCES.

[1] W. L. Chow, "Abelian varieties over function-fields," *Transactions of the American Mathematical Society*, vol. 78 (1955), pp. 253-275.

[2] S. Lang, "Abelian varieties over finite fields," *Proceedings of the National Academy of Sciences*, vol. 41 (1955), pp. 174-176.

[3] T. Matsusaka, "A note on my paper 'Some theorems on abelian varieties'," *Nat. Sc. Rep. Ochanomizu Univ.*, vol. 5 (1954), pp. 21-23.

[4] A. Weil, "On algebraic groups of transformations," *American Journal of Mathematics.*, vol. 77 (1955), pp. 355-391.

[5] O. Zariski, "Theory and applications of holomorphic functions on algebraic varieties over arbitrary groundfields," *Memoirs of the American Mathematical Society*, no. 5 (1951), pp. 1-90.

ON THE REGULARITY REGIONS OF THE SOLUTIONS OF THE PARTIAL DIFFERENTIAL EQUATIONS OF CAUCHY-KOWALEWSKY.*

By AUREL WINTNER.

1. Let $F = F(z, w, s, v)$ be an analytic functions which is regular in a neighborhood of the point $(0, 0, 0, 0)$ of the complex (z, w, s, v)-space and, in a neighborhood of the point $(0, 0)$ of the (z, w)-space, let $s = s(z, w)$ be the (unique, regular) solution of

$$\text{(I)} \qquad s_z = F(z, w, s, s_w), \qquad s(0, w) \equiv 0$$

(Cauchy-Kowalewsky). If $F(z, w, s, v)$ is of the particular form $F(z, s)$, then (I) and its solution $s = s(z, w)$ will simplify to

$$\text{(II)} \qquad ds/dz = F(z, s), \qquad s(0) \equiv 0$$

and $s = s(z)$, respectively. A question concerning an absolute constant in the problem of a lower bound for the convergence radius of the power series $s(z)$ in the special (ordinary) case (II) of the (partial) differential equation (I) was dealt with in [12]. The corresponding question for associated radii of regularity of the solution $s(z, w)$ of the general partial differential equation (I) is much more complex. The only absolute lower bounds known for the associated radii of the double power series $s(z, w)$ are those obtained by Perron [9]. His result will be improved below to some extent, and so as to imply an affirmative answer to one of the questions of function-theoretical interest (concerning *small* functions F of the four variables z, w, s, v) but the problem of the "best absolute constants" remains unsolved.

It appears therefore natural that what should be settled first is the problem of the "best" associated radii of the solution $s(z, w)$ of the initial value problem of a quasi-linear partial differential equation, say of

$$\text{(III)} \qquad s_z = f(z, w, s) + g(z, w, s) s_w, \qquad s(0, w) \equiv 0,$$

with two functions f, g of (z, w, s) which are regular in a neighborhood of

* Received February 29, 1956.

the point $(z, w, s) = (0, 0, 0)$. In fact, the problem for (III) is intermediary between the solved problem for (II) (where $g = 0$) and the unsolved problem for (I) (where F can be more general than $F(z, w, s, v) = f + gv$; $v = s_w$).

2. In order to deal with problems of "best associated radii" in the analytic problem, it seemed to be natural to consult first the literature of the corresponding problem in the real domain. But what the literature contains in this regard proved to be of no avail. It is true true Kamke's book of 1930 contains a theorem which implies, among other things, the following assertion (Satz 4, § 173, pp. 335-336 in [3]): If $f(x, y, u)$ and $g(x, y, u)$ are real-valued, uniformly continuous functions on a parallelepipedon

$$\text{(IV)} \qquad 0 \leqq x < a, \qquad -b < y < b, \qquad -c < u < c$$

and if both functions have, with respect to y and u, continuous first derivatives satisfying

$$\text{(V)} \qquad |d(x, y, u)| \leqq A = \text{const. on (IV), where } d = f_y, f_u, g_y,$$

then there exists an absolute constant $\Theta > 0$ (in fact, some

$$\text{(VI)} \qquad \Theta \geqq \tfrac{1}{2} \log 3;$$

cf. [3], p. 336, footnote) in such a way that the initial value problem

$$\text{(VII)} \qquad u_x = f(x, y, u) u_y + g(x, y, u), \qquad u(0, y) \equiv 0,$$

which is the analogue of (III) in the real field, has on the rectangle

$$\text{(VIII)} \qquad 0 \leqq x < \min(a, \Theta/A), \qquad -b < y < b$$

a (unique) continuously differentiable solution $u = u(x, y)$. But Kamke's proof is erroneous (and, as a matter of fact, the assertion is false for *any* absolute $\Theta > 0$). The trouble is at the very beginning of the proof ([3], p. 336), where, on the one hand, $f(x, y, u)$ and $g(x, y, u)$ are extended from the parallelepipedon (IV), where $0 < b < \infty$, to the infinite slab which is the case $b = \infty$ of (IV) and, on the other hand, the tacit (but erroneous) assumption is made that, when the point (x, y) is in the rectangle (VIII), the point (x, y, u) belonging to the solution $u = u(x, y)$ will stay in the parallelepipedon (IV) belonging to the *given* $b < \infty$. The mistake was pointed out by Kamke himself in a *Nachtrag* to the second edition (1944) of his book (1930) and the result is stated correctly in [4], p. 41, by assuming that $b = \infty$ in (IV) (but the last footnote on p. 42 of [4] is misleading, since the coefficient $f(x, y, u)$ of (VII) above is not allowed to

contain u in [9]). Under the assumption $b=\infty$, following a suggestion of Ważewski (cf. [8], p. 2), Perausówna [8] determined the best value of the absolute constant Θ, by showing that (VI) can be improved to $\Theta \geqq 1$ and that the assumption $\Theta < 1$ is disproved by a simple example; so that $\Theta = 1$ (concerning quasi-linear *systems*, cf. [11]).

It is easy to see that Kamke's proof (with $b=\infty$) and Perausówna's improvement of it (again with $b=\infty$) apply in the complex field also. But what then results applies only to a very particular case of the problem of (III). In fact, Liouville's theorem shows that the analogues of (V) and of the case $b=\infty$ of (IV) in the complex field compel the coefficients of (III) to be of the form

$$f(z,w,s)=\alpha(z)w+\beta(z)s, \qquad g(z,w,s)=\gamma(z)w+\delta(z)s.$$

3. In what follows, the problem of "best associated radii" of (III) will be settled (Theorem 1) by an adaptation of the method used by Perron [10], pp. 557-562, in the particular case $f(x,y,u)=f(x,y)$ of the real equation (VII). By using a device contained already in Kowalewsky's thesis [5] (where the case of her § II is reduced to the case of her § I), the result on the quasi-linear problem (III) will lead to results on the general problem (I) also (Section 7). What thus results for the general case (I) is substantially finer than are the known estimates of the associate radii (even though one cannot readily speak of "best" constants in the general case).

The method, being that of the successive approximations, is such as to supply for (I) a result (Theorem 3) which does not concern the absolute constants but is of independent interest.

4. The following theorem (which, in view of Section 3, contains the central fact of the theory) will be proved first.

THEOREM 1. *Let a, b, c, L, M be five positive numbers, z, w, s three complex variables, finally*

$$(1) \qquad f(z,w,s), \qquad g(z,w,s)$$

two functions which are regular on the (z,w,s)-domain

$$(2) \qquad |z|<a, \qquad |w|+L|z|<b, \qquad |s|<c$$

and satisfy the inequalities

$$(3) \qquad |f|<L, \quad |g|<M \text{ on } (2).$$

Then that solution

$$s = s(z, w) \tag{4}$$

(Cauchy-Kowalewsky) of the quasi-linear partial differential equation

$$s_z = f(z, w, s) s_w + g(z, w, s) \tag{5}$$

which belongs to the initial condition

$$s(0, w) \equiv 0 \tag{6}$$

is regular on the (z, w)-domain

$$|z| < \min(a, c/M), \qquad |w| + L|z| < b \tag{7}$$

(*at least*; but the domain (7) cannot be improved in terms of absolute constants); *moreover, the solution* (4) *is subject to the inequality*

$$|s(z, w)| < c \text{ on } (7). \tag{8}$$

The proof of this theorem will be reduced to a particular case of it:

LEMMA. *Theorem* 1 *is true if f (though not necessarily g) is independent of s, that is, if $f = f(z, w)$.*

If x, w, s and f, g are *real*, then, under appropriate conditions of differentiability (but without the assumption of analyticity, hence, without assuming the applicability of the Cauchy-Kowalewsky theorem), the preceding Lemma is contained in Satz 1 of Perron [10], p. 557, which, in the real field, even generalizes the case $f(z, w, s) = f(z, w)$ of (5) to certain *systems* of partial differential equations. Perron's proof (1927), which was the starting point of *all* the recent developments concerning quasi-linear hyperbolic systems of partial differential equations in two independent variables (for a list of references, cf. [6], pp. 257-258), is based on an application of the process of sucessive approximations. Correspondingly, a glance at Perron's proof shows that it remains valid in the complex field, in the form stated by the Lemma above. As a matter of fact, the proof is now shorter than in Perron's non-analytic case, since only the first, and comparatively easier, half of the proof (pp. 557-559, ending with formula (38)) is needed; the second half of the proof (pp. 559-562), that dealing with the convergence of the *derivatives,* now becomes superfluous, since the uniform convergence of the functions implies that of the derivatives in the analytic case.

5. *Proof of Theorem* 1. For every positive integer n, consider the solution, say

$$(4_n) \qquad s^n = s^n(z, w)$$

(where n is not an exponent), of the partial differential equation

$$(5_n) \qquad s_z^n = f^n(z, w) s_w^n + g(z, w, s^n)$$

and of the initial condition

$$(6_n) \qquad s^n(0, w) \equiv 0,$$

where, if f and g are the two functions given in Theorem 1, the g of (5_n) is the given g, whereas the f^n of (5_n) is defined as follows:

$$(9_n) \qquad f^n(z, w) = f(z, w, s^{n-1}(z, w)).$$

This is an inductive definition, which can be carried out by starting at $n = 0$ with

$$(10) \qquad s^0 = s^0(z, w) \equiv 0.$$

For then, if the (local) theorem of Cauchy-Kowalewsky is applied to (5_n)-(6_n), with (9_n) first for $n = 1$, then for $n = 2, \cdots$, it is clear that each of the functions (4_n) is defined, as a unique regular function, on some neighborhood of the point $(0, 0)$ of the complex (z, w)-space. What fails to follow in this manner is that this neighborhood of $(0, 0)$ can be chosen independent of n. But it will now be concluded from the Lemma that the (z, w)-domain specified by (7), a domain which is independent of n, is such a neighborhood for every n, and that

$$(8_n) \qquad |s^n(z, w)| < c \text{ on } (7).$$

Suppose that, for a fixed n, the function element $s^{n-1}(z, w)$ is known to be regular on (7), and that (8_{n-1}) is true for this n. It will be sufficient to show that both of these assumptions remain valid if $n - 1$ is replaced by n. In view of (10), the induction hypotheses are satisfied if $n - 1 = 0$. If $n - 1$ is fixed (and positive), then the regularity of $s^{n-1}(z, w)$ on (7) and the assumption (8_{n-1}) imply that, since the functions (1) are regular on (2) and satisfy (3), it follows from (9_n) that both functions

$$(1_n) \qquad f^n(z, w), \qquad g(z, w, s)$$

(the first of which is independent of s) are regular on (2) and satisfy

$$(3_n) \qquad |f^n| < L, \ |g| < M \text{ on } (2).$$

Hence, if the Lemma is applied (5_n)-(6_n) (for the fixed n), it follows that the function (4_n) is regular on (7) and satisfies (8_n). This completes the induction.

Next, there exists, in the domain (7), about the origin of the (z, w)-space *a sufficiently small domain*, say

$$(11) \qquad |z| < z_0, \qquad |w| < w_0,$$

on which the sequence of functions

$$(12) \qquad s_1(z, w), \cdots, s_n(z, w), \cdots$$

is convergent. In the real field, this was proved in a paper of Hartman and myself [2], pp. 862-863, by a refinement of the estimates used by Perron [10] in his case, the case of the Lemma (cf. comments above). In the complex field, the proof is exactly the same as in the real field and will therefore be omitted (just as the proof of the Lemma could be omitted above). But as italicized before (11), both radii z_0, w_0 must be chosen sufficiently small before the estimates of [2] assure the convergence of (12) on (11). In fact, the values of the two positive constants z_0, w_0 had to be subjected in [2], p. 862, to quite a number of inequalities (involving the five positive constants occurring in (2)-(3) above).

This could not be avoided in [2] and, correspondingly, it is not avoided in the present case of the complex field. But the "smallness" of the convergence domain (11) of (12) can, in the present case, be disposed of by an appeal to the oldest theorem (Stieltjes) in the theory of normal families (of functions of two variables, to be sure). In fact, since the functions (4_n) are regular on (7) and, according to (8_n), are uniformly bounded on (7), the convergence of the sequence (12) on a "small" dicylinder (11) implies the convergence of (12) on the entire domain (7), and also the uniformity of the convergence on every (z, w)-domain, say D, the closure of which is contained in (7).

Let (4) denote the limit function of (12) on (7). Then $s(z, w)$ is regular on (7). That the limit function is a solution of (5) (even on a sufficiently small (z, w)-domain), is quite an issue in the real field (cf. the end of the proof in [2], pp. 863-864), since the uniform convergence of the derivatives of the functions (12) must also be assured. There is no such issue in the present case, since the uniform convergence of (12) to $s(z, w)$ on every fixed domain D, defined above, implies that the sequences

$$(13) \qquad s_{1z}(z, w), \cdots, s_{nz}(z, w), \cdots \text{ and } s_{1w}(z, w), \cdots, s_{nz}(z, w), \cdots$$

tend to the respective derivatives, $s_z(z,w)$ and $s_w(z,w)$, of $s(z,w)$, uniformly on every fixed domain D. In particular, the three sequences (12), (13) tend to the three functions $s(z,w)$, $s_z(z,w)$, $s_w(z,w)$ at every point (z,w) of the domain (7). Since (4_n) is a solution of (5_n) on (7), it follows that (4) is a solution of (5) on (7). Finally, (6) follows from (6_n), and (8) from (8_n). This completes the proof of the italicized part of Theorem 1.

There remains to be ascertained the parenthetical assertion made after (7). To this end, choose $g(z,w,s)$ to be a function of s alone, and let $f(z,w,s) \equiv 0$. Then (2), (3), (4) and (5), (6) reduce to

$$|s| < c, \quad |g(s)| < M, \quad s = s(z), \quad ds/dz = f(s), \quad s(0) = 0$$

respectively, and (7) can be simplified to the circle $|z| < b/M$. But as shown in [12], there does not exist any *absolute* constant $\epsilon > 0$ in terms of which the circle $|z| < b/M$ could be improved to $|z| < (1+\epsilon)b/M$. For the case in which (5) does not reduce to an ordinary differential equation, cf. the remark to be made in Section 6 concerning the propagation of singularities (characteristics).

6. In contrast to the method of [1], that of [2] consist of successive approximations.* It is the latter method which was used above in order to obtain "sharp" domains of existence in the analytic case, by adapting the considerations of [14].

The propagation of singularities along characteristics, and the nature of the inequalities of Haar (cf. e.g., [4], p. 34 and pp. 119-120) in which this propagation finds it explicit formulation, explain the role played in Theorem 1 by the second of the inequalities (2), the constant L in (2) and (7) being the same as in (3). Correspondingly, it is in the very nature of the problem that something must be lost if the (z,w)-domains considered in assumption (2) and in assertion (7) of Theorem 1 are replaced by (z,w)-dicylinders. The resulting weaker form of Theorem 1 (a weakened

* Since certain recent publications (of none of the authors concerned) make misleading or erroneous statements on a "rediscovery" of the results of Douglis [1] in the last chapter of our paper [2] (a paper the rest of which contains other results also), it should be mentioned here that [1] and [2] were written independently, that the date of receipt of the manuscript of [2] (see p. 834) was December 21, 1951 and that [2] appeared in the October, 1952, issue of the *American Journal of Mathematics*, whereas [1] seems to have appeared sometimes during the summer of 1952 (the exact date and the date of receipt of the manuscript of [1] are not available); so that neither way is it possible to speak of any "rediscovery."

form which cannot, however, be strengthened in terms of (z,w)-dicylinders) is as follows:

THEOREM 2. *Let $f(z,w,s)$, $g(z,w,s)$ be a pair of functions which are regular and bounded, say*

$$|f|<L, \qquad |g|<M, \tag{14}$$

on a tricylinder, say

$$|z|<a, \qquad |w|<b, \qquad |s|<c, \tag{15}$$

about the point $(0,0,0)$ of the space of the complex variables z,w,s. Starting with five such positive numbers L, M, a, b, c, retain the values of L, M and b but replace (if necessary) the given a by a smaller a, and c by a smaller c, so as to satisfy the inequalities

$$0<a<b/L, \qquad 0<c<LM \tag{16}$$

(needless to say, (14) remains true on (15) if the values of a and c, given in (15), are diminished). *Then the solution $s=s(z,w)$ of the quasi-linear initial problem*

$$s_z=f(z,w,s)s_w+g(z,w,s), \qquad s(0,w)\equiv 0 \tag{17}$$

(Cauchy-Kowalewsky) is regular on the dicylinder

$$|z|<\min(a,c/M), \qquad |w|<b-\min(a,c/M) \tag{18}$$

(at least), and

$$|s(z,w)|<c \text{ on } (18). \tag{19}$$

The same is true if $f(z,w,s)$ and $g(z,w,s)$ instead of being regular and subject to (14) on the tricylinder (15), are regular and subject to (14) only on the tricylinder

$$|z|<a, \qquad |w|<b-La, \qquad |s|<c. \tag{20}$$

This is clear from Theorem 1. For, on the one hand, the assumptions (16) assure that (20) and (18) are domains and, on the other hand, the domain (2) is a subset of the tricylinder (20) and the dicylinder (18) is a subset of the domain (7).

7. Instead of the particular case of quasi-linearity, consider now any Cauchy-Kowalewsky equation

$$(21) \qquad u_z = F(z, w, u, u_w)$$

with the initial condition

$$(22) \qquad u(0, w) \equiv 0$$

for the solution $u = u(z, w)$ and with an $F = F(z, w, u, v)$ which is given as a regular function of its four variables on a neighborhood of the origin of the complex (z, w, u, v)-space, say on the domain

$$(23) \qquad |z| < a, \qquad |w| < b, \qquad |u| < c, \qquad |v| < d.$$

If a, b, c, d are replaced (when necessary) by somewhat smaller positive numbers a, b, c, d, then F is bounded on (23), say

$$(24) \qquad |F(z, w, u, v)| < C \text{ on } (23).$$

It follows from (24), and from Cauchy's inequalities for the coefficients of the power series of F, that (21) and (22) are majorized by (21*) and (22) if (21*) denotes the equations which results from (21) if F is replaced by F^*, where

$$(25) \qquad F^*(z, w, u, v) = C/\{(1 - z/a)(1 - w/b)(1 - u/c)(1 - v/d)\}$$

on (23). But consider first (21) itself and apply to it a formal device (used already in Kowalewsky's thesis [5]), in a way which reduces (21)-(22) to the quasi-linear case (but in a less drastic form as in [5]), as follows:

First, if $v = v(z, w)$ denotes the partial derivative $u_w = u_w(z, w)$, then

$$(26) \qquad u(0, w) \equiv 0, \qquad v(0, w) \equiv 0,$$

by (22), and partial differentiation of (21) with respect to w shows that

$$(27) \qquad \begin{aligned} u_z &= F(z, w, u, v) + 0 \, . \, u_w, \\ v_z &= H(z, w, u, v) + F_v(z, w, u, v) \, . \, v_w, \end{aligned}$$

where

$$(28) \qquad H(z, w, u, v) = F_w(z, w, u, v) + vF_u(z, w, u, v),$$

since $u_{zw} = u_{wz}$. But the comparison of like powers of z, w shows that there is a *unique* pair of (formal) power series u, v in (z, w) satisfying (27) and (26), and that there is a *unique* (formal) power series in (z, w) satisfying (21) and (22). Consequently, (27) and (26) together are not only necessary but sufficiently as well for (21) and (22) together.

where $\alpha = g_w$, $\beta = g_u - f_z$, $\gamma = -f_u$; so that, if a solution $(w, u) = (w(z), u(z))$ of (37) is known, then (38) reduces to a *Riccati equation*

(38 bis) $$dv/dz = a(z) + b(z)v + c(z)v^2$$

for $v = v(z)$. Finally, if $f(z, w, u)$ is assumed to have Perron's particular form, that is, if $f = f(z, w)$ (cf. the Lemma of Section 4), then the binary system (37) *splits* into two equations of first order:

(37_1) $\quad dw/dz = f(z, w),$ $\qquad$ (37_2) $\quad du/dz = g(z, w, u).$

In fact, if a solution $w = w(z)$ of (37_1) is known, then (37_2) becomes of the form

$(37_2\,\text{bis})$ $$du/dz = h(z, u),$$

and is therefore, like (37_1), a differential equation for a single function.

10. The following considerations deal, like Section 7, with the Cauchy-Kowalewsky majorant (21*)-(22) of (21)-(22); cf. (25) and (24). But the method will be more direct, and the result more explicit, than in Section 7.

Let the notation be so chosen that all four radii occurring in (23) become 1. Then (24) and (25) show that the Cauchy-Kowalewsky majorant (21*) of (21) becomes

(40) $$\phi_z = C/\{(1-z)(1-w)(1-\phi)(1-\phi_w)\},$$

where C is a positive constant. Let $\phi(z, w)$ denote that solution of (40) which is determined by the initial condition

(41) $$\phi(0, w) \equiv 0;$$

cf. (21) and (22), where $u = \phi$.

According to Perron [9], p. 158, the function $\phi(z, w) = \phi_C(z, w)$ is regular in the dicylinder

(42) $$|z| < (1-r)^2/(1+8C), \qquad |w| < r,$$

at least, if r is any number on the interval $0 < r < 1$. It is however natural to expect that, in view of (40), the location of the singularities will be such that *every point* (z, w) *of the dicylinder* ($|z| < 1$, $|w| < 1$) *becomes a regular point of* $\phi_C(z, w)$ *as* $C \to 0$. But (42) it too rough to prove this. It will be shown, however, that (42) can be improved to

(43) $$|z| < 1 - \exp(-\tfrac{1}{8}(1-r)^2/C), \qquad |w| < r.$$

This result (the proof of which will depend on an application of (47) below) proves the italicized desideratum, since if $r = 1 - \epsilon < 1$ is fixed and $C = C_\epsilon$ is large enough, then the dicylinder (43) contains the dicylinder $(|z| < 1 - \epsilon, |w| < 1 - \epsilon)$. It will be seen in what follows that the *structure of* (43) *is "correct,"* in the sense that (43) is close to the ultimate truth.

In order to formulate this situation precisely, let C be fixed in (40) and let r be any positive number corresponding to which there exists a positive R having the property that the solution $\phi(z, w)$ of (40) and (41) is regular in the dicylinder $(|z| < R, |w| < r)$. By $R = R(r)$ will be meant the greatest such R (so that $(R(r), r)$ actually is a pair of associated radii, a pair to which Hartogs's convexity theorem applies).

THEOREM 4. *If $R(r) = R_C(r)$ denotes the z-radius associated to a w-radius $r > 0$ in the solution $\phi = \phi(z, w) = \phi_C(z, w)$ of (40) and (41), then, on the one hand*

$$R(r) = 0 \text{ if } r = 1 \text{ (hence, if } r > 1) \tag{44}$$

and, on the other hand,

$$R(r) = 1 - \exp(-(1 - r)S(r)/C) \text{ if } 0 < r < 1, \tag{45}$$

where $S(r)$ is a function satisfying the inequalities

$$\tfrac{1}{8}(1 - r) \leqq S(r) \leqq \tfrac{1}{2} \qquad (0 < r < 1) \tag{46}$$

(cf. also (49) below, where $\gamma > 0$).

Both of the equality signs can be excluded in (46), since it will be clear from the proof of (46) that neither of the constants $\frac{1}{2}$, $\frac{1}{8}$ can be "best"; cf. also (49) below. Note that C occurs explicitly in (45) but not in (46), even though it is not clear that $S(r)$ is not a function, $S_C(r)$, of C also. Since (45) and the second of the inequalities (46) imply that $R(r) \to 0$ as $r \to 1$, they imply (44).

11. First, since the coefficients of all the power series involved are positive (or 0), it is clear that a minorant of (40) and (41) results if (40) is retained and (40) is replaced by the partial differential equation, say (40 bis), which results if the fourth factor, $(1 - \phi_w)$, of the denominator is omitted in (40). But (40 bis) can be written in the form

$$\phi_z = \mu / \{(1 - z)(1 - \phi)\}, \qquad \mu = C/(1 - w), \tag{40 bis}$$

where μ is a parameter, and so (40 bis) is an (ordinary, rather than a partial) differential equation which the substitution

$$(47) \qquad t = -\log(1-z) \qquad (|z|<1,\ \log 1 = 0)$$

transforms into $d\phi/dt = \mu(1-\phi)$. The solution $\phi = \phi(t)$ is given by $(1-\phi)^2 = -2\mu t + \text{const.}$, where const. $= 0$ by virtue of (41) and (47). It follows therefore from (47) that the solution $\phi = \phi(z; w)$ of (40 bis) becomes singular at the z-value $z_0 = z_0(\mu)$ determined by the equation $(1-z_0)^{2\mu} = 1/e$. Since $\mu = C/(1-w)$, this equation leads to

$$(48) \qquad z_0 = 1 - \exp(-\tfrac{1}{2}(1-r)/C)$$

if $w = r$, where $0 < r < 1$.

Since (40) is a majorant of (40 bis) (with reference to the common initial condition (41)), it follows that the solution $\phi(z, w)$ of (40) and (41) cannot be regular on the closure of the dicylinder $(|z| < z_0, |w| < r)$ if z_0 is the value determined by (48), where $0 < r < 1$, hence $0 < z_0 < 1$. In view of the definition of $R(r)$, this proves the second of the inequalities (46) for the function $S(r)$ defined by (45). Incidentally, it is easy to see that the $\frac{1}{2}$ in (46) can be improved to $\frac{1}{3}$ if the factor $1/(1-\phi_w)$ in (41) is relaxed only to $1/(1-\phi)$ and not, as in the replacement of (40) by (40 bis), to $1/(1-0) = 1$. But this is unimportant, since what is missing in (46) is the improvement of the constant upper bound for $S(r)$ to an upper bound which depends on r (cf. Section 11 below), in the same way as the lower bound supplied by (46). In view of (45), this lower bound for $S(r)$ is equivalent to the result (43) of [13].

Concerning an r-dependent improvement of the upper bound in (46), I find, by using minorants of (40) which are more elaborate than (40 bis), that

$$(49) \qquad S(r) = O(1-r)^{\gamma} \text{ as } r \to 1$$

holds for some constant $\gamma > 0$ (for instance, for $\gamma = \frac{2}{3}$) and the computations seem to indicate that (49) holds for any $\gamma < 1$. If (49) should be true for $\gamma = 1$ also, then, writing (45) in the form

$$(50) \qquad R(r) = 1 - \exp(-(1-r)^2 T(r)/C), \text{ where } 0 < r < 1,$$

one could conclude from the lower bound supplied by (46) that

$$(50?) \qquad 0 < \operatorname*{fin\,inf}_{0<r<1} T(r) \leqq \operatorname*{fin\,sup}_{0<r<1} T(r) < \infty;$$

so that the problem of the pair of the associated radii $(R(r), r)$ of $\phi(z, w)$

would become refined to the determination of the two positive values occurring in (50?).

12. There remains to be proved the assertion of Theorem 4 concerning the first of the inequalities (46). In view of (45), that assertion can be formulated as follows: If r is any number on the interval $0 < r < 1$, then the solution $\phi = \phi(z, w)$ of (40) and (41) is regular on the dicylinder

$$(51) \qquad |z| < 1 - \exp(-(1-r)^2/8C), \qquad |w| < r$$

(and, therefore, on the union of the r-family (51) of dicylinders).

Perron [9] obtained his domain (42) by (retaining (41) but) *first* majorizing (40) by

$$(52) \qquad \phi_z = C/\{(1-z)(1-w)(1-(1-w)^{-1}\phi - \phi_w)\}$$

and *then* (52) by

$$(53) \qquad \phi_z = C/\{(1-z)(1-[1-z]^{-2}w)(1-[1-w]^{-1}\phi - \phi_w)\}.$$

But it will be seen that the device, applied by Perron to an "explicit" integration of (53), can be adjusted so as to work for (52) itself, rather than just for the weakened form, (53), of (52). This will be made possible by an application to (52) of the mapping (47). The result will be that the solution $\phi(z, w)$ of (52) and (41) is regular on any dicylinder (51), where $0 < r < 1$.

First, (47) transforms (52) into

$$(54) \qquad (1-w)\phi_t = C/\{1-(1-w)^{-1}\phi - \phi_w\}$$

but leaves (41) unaltered, since $z = 0$ corresponds to $t = 0$ in (47). Next, replace $\phi = \phi(t, w)$ by $\psi = \psi(t, w)$, where

$$(55) \qquad \psi = \phi/(1-w).$$

Then it turns out that, just as in Perron's case (cf. [9], pp. 157-158), the function $\psi = \psi(t, w)$ (determined, near $(t, w) = (0, 0)$, by (53), (54) and (41)), is a function $\psi = \psi(u)$ of a *single* variable u, to be defined by

$$(56) \qquad u = t/(1-w)^2.$$

In fact, a straightforward calculation shows that, by virtue of (55) and (56), the *partial* differential equation (54) is identical with the *ordinary* differential equation

$$(57) \qquad \psi' = C/(1-2u\psi'),$$

where $\psi = \psi(u)$ and $\psi' = d\psi/du$, whilst (41) goes over into the initial condition $\psi(0) = 0$.

It is clear that (57) is a quadratic equation for $\psi' = \psi'(u)$, and that its root ψ' which remains regular at $u = 0$ is

$$(58) \qquad \psi' = (1 - (1 - 8Cu)^{\frac{1}{2}})/4u.$$

The function (58) of u is regular on the circle

$$(59) \qquad |u| < (8C)^{-1}$$

of the u-plane. Since $\psi = \psi(u)$ follows from (58) (and from the initial condition $\psi(0) = 0$) by a quadrature, $\psi(u)$ itself must be regular on the circle (59). This completes the proof. For, on the one hand, $\phi(z, w)$ is identical with $\psi(u)$ by virtue of (47), (55) and (56), and, on the other hand, it is readily seen from (47) and (56) that the u-circle (59) corresponds to a (z, w)-domain which contains the dicylinder (51) belonging to an arbitrary value of r, where $0 < r < 1$.

THE JOHNS HOPKINS UNIVERSITY.

REFERENCES.

[1] A. Douglis, "Existence theorems for hyperbolic systems," *Communications of Pure and Applied Mathematics*, vol. 5 (1952), pp. 119-159.

[2] P. Hartman and A. Wintner, "On hyperbolic partial differential equations," *American Journal of Mathematics*, vol. 74 (1952), pp. 834-864.

[3] E. Kamke, *Differentialgleichungen reeller Funktionen*, Leipzig, 1930.

[4] ———, *Differentialgleichungen*, Lösungmethoden und Lösungen, vol. II, Leipzig, 1944.

[5] S. von Kowalewsky, "Zur Theorie der partiellen Differentialgleichungen," *Crelle's Journal*, vol. 80 (1875), pp. 1-32.

[6] P. D. Lax, "Nonlinear hyperbolic equations," *Communications of Pure and Applied Mathematics*, vol. 6 (1953), pp. 231-258.

[7] P. Painlevé, "Sur le calcul des intégrales d'un système différentiel par la méthode de Cauchy-Lipschitz," *Bulletin de la Société Mathématique de France*, vol. 27 (1899), pp. 149-152.

[8] J. Perausówna, "Sur le domaine d'existence des intégrales de l'équation $p + f(x, y, z)q = g(x, y, z)$," *Annales de la Societé Polonais de Mathématique*, vol. 12 (1933), pp. 1-5.

[9] O. Perron, "Ein neuer Bewis des Fundamentalsatzes in der Theorie der partiellen Differentialgleichungen erster Ordnung," *Mathematische Zeitschrift*, vol. 5 (1919), pp. 154-160.

[10] ———, "Ueber Existenz und Nichtexistenz von Integralen partieller Differentialgleichungenssytème im reellen Gebiet," *ibid.*, vol. 27 (1928), pp. 549-564.

[11] T. Ważewski, "Sur le domaine d'existence des intégrales de l'équations aux derivées partielles du premier ordre linéaire," *Annuales de la Société Polonaise de Mathématique*, vol. 12 (1933), pp. 6-15.

[12] A. Wintner, "On the exact value of the bound for the regularity of solutions of ordinary differential equations," *American Journal of Mathematics*, vol. 57 (1935), pp. 539-540.

[13] ———, "On the method of successive approximations in initial value problems," to appear in the *Annali di Matematica*, 1956.

ON CERTAIN ABSOLUTE CONSTANTS CONCERNING ANALYTIC DIFFERENTIAL EQUATIONS.*

By AUREL WINTNER.

PART I.

This note on *ordinary* differential equations, though it does not presuppose the paper on *partial* differential equations which precedes it in this volume (pp. 525-541), can be thought of as an appendix to that paper, since it contains (among other things) a justification of the choice of the region of analyticity in Theorem 1 *loc. cit.*

Consider the solution $w = w(z)$ of the differential equation

$$dw/dz = f(z, w) \tag{1}$$

with the initial condition

$$w(0) = 0 \tag{2}$$

and under the following assumptions: $f(z, w)$ is regular on the dicylinder

$$|z| < a, \qquad |w| < b \tag{3}$$

and is bounded there, say

$$|f(z, w)| < M \text{ on } (3). \tag{4}$$

Then a classical result states that the method of successive approximations supplies the circle

$$|z| < \min(a, b/M) \tag{5}$$

on which the solution $w(z)$ of (1) and (2) is sure to be regular.

For a long time, and by methods usually distinct from that of the successive approximations, various efforts have been made in order to improve on the radius of the circle (5) of assured regularity for $w(z)$; cf. M. Müller's report in [6], pp. 169-172. But it was shown in [8] that those efforts

* Received February 29, 1956.

could not possibly succeed, since, if $f(z, w)$ is independent of z (so that (1) and (5) reduce to

(6) $$dw/dz = f(w)$$

and $|z| < b/M$ respectively), then, corresponding to every $\epsilon > 0$, it is possible to exhibit an $f(w) = f_\epsilon(w)$ having the property that the solution $w(z)$ of (6) and (2) will possess a singularity within the circle $|z| < (1+\epsilon)b/M$.

In this sense, (5) is the best possible result. But as mentioned in [8], this result depends on the assumption that no supplementary information, such as a Lipschitz constant, is added to the estimate (4). The purpose of the following considerations is to fill in the resulting gap.

Let $f(z, w)$ be regular on (3) but, instead of assuming (4), suppose only that $f(z, 0)$ is bounded, say

(7) $$|f(z, 0)| \leqq N \text{ on } |z| < a,$$

but suppose also that the partial derivative of $f(z, w)$ with respect to w is bounded, say

(8) $$|\partial f(z, w)/\partial w| < L \text{ on } (3).$$

Although the field is complex, it is readily seen that (8) is equivalent to Lipschitz's condition

(8 bis) $$|f(z, w') - f(z, w'')| < L\,|w' - w''|,$$

where (z, w'), (z, w'') are any two points (z, w) on (3) with distinct w but common z. Since (8) is equivalent to (8 bis), it follows from a remark of Painlevé [7] on a result of E. Lindelöf [4], p. 123 (concerning successive approximations), that the solution $w(z)$ of (1) and (2) is regular on the circle

(9) $$|z| < \min(a,\ L^{-1}\log(1 + bL/N))$$

(at least). Before Lindelöf (but not with the method of successive approximation), and in the real field, the domain (9) was found by Lipschitz himself ([5], pp. 509-514). Cf. also a paper of O. Hölder [2].

The choice of the radius of the circle (9), in contrast to that of (5), seems to be artificial (except, perhaps, if an appeal is made to the inequality of Haar [1] in the theory of characteristics). But it turns out that (9), like (5), is the best possible result of its own kind. By this is meant that, if either the a or the $L^{-1}\log$ in (9) is multiplied by $1+\epsilon$, where $\epsilon > 0$ is arbitrarily small, then there results a z-circle within which the solution $w(z)$ of (1) and (2) will acquire a singularity, if $f(z, w) = f_\epsilon(z, w)$ is suitably chosen.

This is clear for the replacement of a by $(1+\epsilon)a$ in (9). In fact, if $f(z,w)$ is independent of w, then (1) and (9) reduce to $dw/dz=f(z)$ and $|z|<a$ respectively, and so the solution $w(z)$ of (1) and (2) will become singular on the boundary $|z|=a$ whenever $f(z)$ does. Consequently, it is sufficient to consider the case complementary to the case $f(z,w)=f(z)$, that is, the case $f(z,w)=f(w)$.

Then, if $b=1$ without loss of generality, (1) and (9) reduce to (6) and

$$(10) \qquad |z|<L^{-1}\log(1+L/|f(0)|)$$

respectively, since the N in (7) can be chosen to be the value of $|f(w)|$ at $w=0$ (this value can be assumed to be distinct from 0, since $w(z)\equiv 0$ is the solution of (6) and (2) if $f(0)=0$). On the other hand, the formulation (8) of Lipschitz's condition (8 bis) reduces to the assumption that

$$(11) \qquad |f'(w)|<L \text{ for } |w|<1,$$

where $f'=df/dw$ and $1=b$. Accordingly the assertion, to be proved, is as follows: If it is only assumed that the function $f(w)$ is regular on the circle $|w|<1$ and that its derivative is restricted by (11), then the solution $w(z)$ of (6) and (2) *can* become singular at $z=z^0$ whenever $|z^0|$ is given as a number exceeding the radius of the circle (10).

Choose $f(0)=1$ and write ϵ instead of L. Then $f(w)$ is a power series of the form

$$(12) \qquad f(w)=1+g(w), \text{ where } g(w)=\sum_{n=1}^{\infty} c_n w^n$$

for $|w|<1$, the circle (10) reduces to

$$(13) \qquad |z|<\epsilon^{-1}\log(1+\epsilon)=1+o(1) \text{ if } \epsilon\to 0,$$

and condition (11) is satisfied (for a fixed ϵ) if

$$(14) \qquad \sum_{n=1}^{\infty} n\,|c_n|<\epsilon,$$

where $c_n=c_n(\epsilon)$. On the other hand, it is clear from (12) that, if ϵ is fixed in $g(w)=g_\epsilon(w)$, then, near $z=0$, the solution $w=w(z)=w_\epsilon(z)$ of (6) and (2) is the (local) inverse of the function $z=z(w)=z_\epsilon(w)$ defined by

$$(15) \qquad z(w)=\int_0^w (1+g(w))^{-1}\,dw, \text{ where } g(0)=0.$$

The only restriction on the coefficients $c_n = c_n(\epsilon)$ of the power series (12) is that they should be "small" in the sense of (14). Hence it is easily realized (cf. a parallel construction in the proof of assertion (β^*) of (ii) below) that, for every $\epsilon > 0$, it is possible to choose the coefficients $c_n = c_n(\epsilon)$ of $g(w) = g_\epsilon(w)$ in accordance with (14) and in such a way that the function $w(z) = w_\epsilon(z)$, defined near $z = 0$ as the inverse of the function (15), becomes singular at some point of a circle $|z| < r_\epsilon$ the radius of which is of the form $r_\epsilon = 1 + o(1)$ as $\epsilon \to 0$. But this is precisely the assertion, since the circle (10) is represented by (13).

PART II.

If $a = 1$, $b = 1$ and $M = 1$ in (3)-(4), then the circle (5) becomes $|z| < 1$ and, in view of the maximum principle. nothing is lost if the sign of equality is allowed in (4). If this is combined with the circumstance that successive approximations $w_0(z) \equiv 1$, $w_1(z), \cdots, w_n(z), \cdots$ which lead to the solution $w(z)$ of (1) and (2) are subject to the inequality $|w_n(z)| < b$ on the circle (5), then there results the following fact (i):

(i) *If $f(z,w)$ is regular, and satisfies the inequality $|f(z,w)| \leqq 1$, on the dicylinder $(|z| < 1, |w| < 1)$, then the solution $w(z)$ of (1) and (2) is regular, and satisfies the inequality $|w(z)| < 1$, on the circle $|z| < 1$.*

For a refinement of (i), cf. [14].

Other applications of known "best contants" pertaining to classes of power series result if (6) is contrasted with

$$dw/dz = f^*(w), \tag{16}$$

where

$$f^*(w) = \sum_{n=0}^{\infty} |c_n| w^n \text{ if } f(w) = \sum_{n=0}^{\infty} c_n w^n. \tag{17}$$

In order to see this, suppose that $f(w)$, hence $f^*(w)$, is regular on the circle $|w| < 1$. Then it is clear from the proof of (i) that the solution $w(z)$ of (16) and (2) is regular on any circle

$$|z| < r/f^*(r), \tag{18}$$

where r is any positive number less than 1 (for a direct proof, cf. Lindelöf's method in [11]).

Suppose that $f(w)$, besides being regular for $|w| < 1$, satisfies the inequality

$$|f(w)| < 1 \text{ for } |w| < 1. \tag{19}$$

Then, according to Bohr (cf. [10], pp. 32-34), the inequality $f^*(r) \leqq 1$ holds if $r = \frac{1}{3}$ but not in general if $r = \frac{1}{3} + \epsilon > \frac{1}{3}$. It follows from the first of these two facts (the second fact, that concerning the non-existence of an absolute constant $\epsilon > 0$, is irrelevant this time) that the radius of the circle (18) becomes not less than $\frac{1}{3}$ at $r = \frac{1}{3}$. Consequently, (19) is sufficient in order that the solution $w(z)$ of (16) and (2) be regular on the circle $|z| < \frac{1}{3}$.

On the other hand, as will be shown elsewhere [16], even the most favorable choice of r (on the range $0 < r < 1$) can bring the radius of (18) arbitrarily close to $\frac{1}{3}$ if only (19) is assumed. One might therefore expect that $\gamma = \frac{1}{3}$ is the greatest absolute constant having the property that the solution $w(z)$ of (16) and (2) is regular on the circle $|z| < \gamma$ whenever (19) holds. This expectation proves, however, to be erroneous. The explanation is that $\frac{1}{3}$ is the best constant by virtue of Cauchy's principle of "majorization," but *not* by virtue of the finer principle of "subordination" (in this regard, cf. the concluding remarks in [15]). In fact, it turns out that $|z| < \frac{1}{3}$ can be improved to $|z| < \gamma$ (but not, of course, to $|z| < 1$), where γ denotes the least positive number r (< 1) for which the function

$$r(1 - r^2)^{-\frac{1}{2}} - \arcsin r \qquad (0 \leqq r < 1) \tag{21}$$

becomes 1 (in (21), both the square root and the arc sin are meant to be positive).

(I) *If $f(w)$ is regular on the circle $|w| < 1$ and satisfies (19), then, while the solution $w(z)$ of (6) and (2) is regular, and satisfies the inequality $|w(z)| < 1$, on the circle $|z| < 1$, the solution $w(z)$ of (16) and (2) is regular, and satisfies the inequality $|w(z)| < 1$, on the circle $|z| < \gamma$, where γ denotes the least positive root of the transcendental equation*

$$\gamma/(1 - \gamma^2)^{\frac{1}{2}} = 1 + \arcsin \gamma \qquad (\arcsin 0 = 0). \tag{22}$$

The first of the two assertions of this theorem (I) is contained in (i). In order to prove the second assertion of (I), note that, according to (17),

$$|f^*(w)|^2 \leqq \sum_{n=0}^{\infty} |c_n|^2 \sum_{n=0}^{\infty} |w|^{2n}, \text{ where } \sum_{n=0}^{\infty} |c_n|^2 < 1,$$

by (19), except when $f(z) = cz$, where $|c| = 1$. Hence, in any case,

$$f^*(r) < \left(\sum_{n=0}^{\infty} r^{2n}\right)^{\frac{1}{2}} = (1 - r^2)^{-\frac{1}{2}} \text{ if } 0 < r < 1, \tag{23}$$

where $f^*(r) \geqq 0$, by (17).

Recourse can now be had to a theorem on "subordination," initiated by Nakano [12] and formulated in a general form in [15]. In fact, it is clear from (23) and from the lemma of [15], p. 1106, that the solution $w(z)$ of (16) and (2) is regular, and satisfies the inequality $|w(z)| < 1$, on the circle $|z| < \gamma$, if γ is any positive number having the following property: If the function $r = r(t)$ $(\geqq 0)$ is defined as the solution of the differential equation

$$(24) \qquad dr/dt = (1 - r^2)^{-\frac{1}{2}}$$

and of the initial condition

$$(25) \qquad r(0) = 0,$$

then $r(t)$ exists, and satisfies the inequality $r(t) < 1$, on the interval $0 \leqq t < \gamma$. But (24) can be solved by the inversion of a quadrature, whence it is seen that the solution $r = r(t)$ of (24) and (25) is the (local) inverse of the function $t(r)$ which results if the function (21) of r is denoted by $t = t(r)$. Finally, it is clear that the value of the function (21) is between 0 and 1 for $0 < r < \gamma$, if γ is the least positive root (< 1) of the equation (22). This completes the proof of (I).

Since the estimate (23) of $f^*(r)$ by $(1 - r^2)^{-\frac{1}{2}}$ excludes the sign of equality (except, perhaps, at $r = 0$), it also follows that *the* γ *of* (I) *can be improved to* $\gamma + \epsilon$, *where* $\epsilon = \epsilon_f > 0$. On the other hand, it remains undecided whether this ϵ can be chosen independent of f (in other words, whether γ is the best absolute constant). In this regard, nothing seems to follow from a direct consideration of the sequence of rational functions which, for quite another purpose, Landau has constructed from the sequence of the partial sums of

$$(23\,\text{bis}) \qquad (1 - W)^{-\frac{1}{2}} = \sum_{n=0}^{\infty} b_n W^n, \text{ where } b_n = 1.3 \cdots (2n - 1)/2.4 \cdots 2n$$

($W = w^2$ in (20); cf. [10], pp. 26-29, and a general theorem of Schur [13], pp. 122-124).

It will now be shown that if the circle $|z| < \gamma$, defined in (I), is replaced by the smaller circle $|z| < \sin \gamma$, then (I) can be transferred to the case in which (16) is generalized to

$$(26) \qquad dw/dz = f^*(z, w),$$

where, corresponding to (17),

$$(27) \qquad f^*(z,w) = \sum_{m=0}^{\infty} \sum_{n=0}^{\infty} |c_{mn}| z^m w^n \text{ if } f(z,w) = \sum_{m=0}^{\infty} \sum_{n=0}^{\infty} c_{mn} z^m w^n$$

(so that $f^*(z,w)$ is regular on a dicylinder (3) whenever $f(z,w)$ is).

(II) *If $f(z,w)$ is regular, and satisfies the inequality $|f(z,w)| < 1$, on the dicylinder $(|z| < 1, |w| < 1)$, then the solution $w(z)$ of* (26) *and* (2) *is regular, and satisfies the inequality $|w(z)| < 1$, on the circle $|z| < \sin\gamma$, where γ (< 1) denotes the least positive root of the transcendental equation* (22).

In view of (i), the content of (II) is that, if (1) is replaced by (26), then the circle $|z| < 1$, supplied by (i), becomes replaced by a circle $|z| < \lambda$, where λ is an absolute constant (< 1) which (II) claims to be not less than $\sin\gamma$. It can readily be shown that, corresponding to the remarks made before (I) on the $\frac{1}{3}$-radius, the simple method of Lindelöf [11] supplies for (26) and (2) an absolute constant λ which is substantially less favorable than the absolute constant supplied by (II). The proof of (II) proceeds as follows:

Clearly, the assumptions of (II) and the second of the relations (27) imply that, corresponding to the proof of (20),

$$\sum_{m=0}^{\infty} \sum_{n=0}^{\infty} |c_{mn}|^2 \leqq 1; \text{ hence } f^*(|z|, |w|) \leqq \Big(\sum_{m=0}^{\infty} \sum_{n=0}^{\infty} |z|^{2m} |w|^{2n}\Big)^{\frac{1}{2}}$$

if $0 < |z| < 1$, $0 < |w| < 1$, by the first of the relations (27). Since the preceding square root is $(1-s^2)^{-\frac{1}{2}}(1-r^2)^{-\frac{1}{2}}$ if $|z| = s < 1$, $|w| = r < 1$, it follows that (26) and (2) are "subordinated" to

$$(28) \qquad dr/ds = (1-s^2)^{-\frac{1}{2}}(1-r^2)^{-\frac{1}{2}}$$

and (25). In fact, if [15] is applied in the same way as in the proof of (I), it follows that the solution $w(z)$ of (26) and (2) is regular, and in absolute value less than 1, on any circle $|z| < \lambda$ the radius of which is a positive number satisfying $\lambda \leqq 1$ and having the following property: The solution $r = r(s)$ $(\geqq 0)$ of (26) and (25) exists, and satisfies the inequality $r(s) < 1$, on the interval $0 \leqq s < \lambda$.

On the other hand, if

$$(29) \qquad t = \arcsin s, \text{ where } 0 \leqq t < \tfrac{1}{2}\pi, \qquad 0 \leqq s < 1,$$

then t is increasing with s (and $s = 0$ corresponds to $t = 0$), and (29) transforms (28) into (24) and leaves (25) unaltered. But the solution $r = r(t)$ of (24) and (25) exists, and satisfies the inequalities $0 \leqq r(t) < 1$, on the

interval $0 \leqq t < \gamma$ (cf. the end of the proof (I)). It follows therefore from (29) that the solution $r = r(s)$ of (28) and (25) exists, and satifies the inequalities $0 \leqq r(s) < 1$, on the interval $0 \leqq s < \lambda$, where $\lambda = \sin \gamma$. In view of the end of the preceding paragraph, this completes the proof of (II).

It is instructive to compare the above proofs and results with those of Cauchy himself (cf., e. g., pp. 6-7 of Bieberbach's book (1953) on the complex function theory of ordinary differential equations). Cauchy assumes that $f(z, w)$ is regular and bounded on a dicylinder (3). It can be assumed that $a = 1$ and $b = 1$ in (3), and that (4) holds for $M = 1$. Then the assumptions on $f(z, w)$ are precisely the same as in (i) or (II). But instead of Parseval's relation and Schwarz's inequality, used in the proof of (II) above, and instead of using the principle of *subordination* (which is the crucial step in the above proof; cf. [15], p. 1107), Cauchy applies his inequalities, $|c_{mn}| \leqq M/a^m b^n$, for the coefficients of (27), which commits him to the following *majorization* of the power series (27):

$$|f(z, w)| \leqq f^*(|z|, |w|) \leqq \sum_{m=0}^{\infty} \sum_{n=0}^{\infty} 1 \cdot |z|^m |w|^n,$$

since $M/a^m b^n = 1$ in the present normalization. Accordingly, not only (1) but also (26) is being majorized by

$$dw/dz = (1 - z)^{-1}(1 - w)^{-1} \tag{30}$$

and, correspondingly, (28) is roughened to

$$dr/ds = (1 - r)^{-1}(1 - s)^{-1} \tag{31}$$

(leaving no possibility for applying "subordination" instead of "majorization").

The substitution $Z = -\log(1 - z)$, where $|z| < 1$ (and $Z = 0$ at $z = 0$), reduces (30) to $dw/dZ = (1 - w)^{-1}$. The solution of the latter differential equation and of (2) is $w = \frac{1}{2} - \frac{1}{2}(1 - 2Z)^{\frac{1}{2}}$, a function which is regular on the circle $|Z| < \frac{1}{2}$ but not at the point $Z = \frac{1}{2}$. Hence it is clear from $z = 1 - e^{-Z}$ that the solution $w(z)$ of (30) and (2) is regular on the circle

$$|z| < \lambda, \text{ where } \lambda = 1 - e^{-\frac{1}{2}}, \tag{32}$$

and becomes singular at the boundary point $z = \lambda$ of the circle (32). It follows that the solution $w(z)$ of (26) and (2) is regular on the circle (32).

But this circle is substantially smaller than the circle, $|z| < \sin\gamma$, supplied by (II).

It may finally be mentioned that the discrepancy between the respective radii, γ and $\sin\gamma$ $(<\gamma)$, supplied by (I) and (II), has an analogue when only Cauchy's majorants are applied. In fact, if (26) is of the particular form (16), then (30) can be replaced by $dw/dz = (1-w)^{-1}$. Since the solution of the latter differential equation and of (2) is $w = \frac{1}{2} - \frac{1}{2}(1-2z)^{\frac{1}{2}}$, it follows that Cauchy's method of majorants improves his circle (32) to $|z| < \frac{1}{2}$ (though not to any circle $|z| < \frac{1}{2} + \epsilon$, where $\epsilon > 0$), if (26) is of the particular form (16).

In the particular case (6) of (1), the information contained in (i) can be completed as follows:

(ii) *On the circle $|w| < 1$, let $f(w)$ be a regular function*

$$f(w) = \sum_{n=0}^{\infty} c_n w^n \tag{33}$$

satisfying (19). *Denote by $w(z) = w_f(z)$ the solution of* (6) *and* (2). *Then*

(α) *the function $w(z)$ is regular not only on the circle $|z| < 1$ but also on some circle $|z| < 1 + \epsilon$, where $\epsilon = \epsilon_f > 0$ (and, except when $f(w) \equiv f(0)$ and $|f(0)| = 1$, not only $|w(z)| < 1$ but also $|w(z)| <$ const. < 1 holds for $|z| < 1$) but*

(β) *the positive number $\epsilon = \epsilon_f$ cannot be chosen independent of f in* (α) *and, what is more,*

(β^*) *even if $f(w)$ is restricted by $c_n \geqq 0$ for every n in* (33) *(that is, even if $f(w) \equiv f^*(w)$), there belongs to every $\epsilon > 0$ some $f(w) = f^\epsilon(w)$ corresponding to which the function $w(z) = w_f(z)$ becomes singular within the circle $|z| < 1 + \epsilon$.*

(α) is a particular case of what was proved in [14] and (β) is the result of [8]. In view of the contrast between (I) and (II), the improvement of (β) to (β^*) is relevant from the point of *any* majorant of (6); in fact, (6) is its own *best majorant* if $f = f^*$. The proof of (β^*) proceeds as follows:

Let $0 < \epsilon < 1$ (eventually, $\epsilon \to 0$) and choose c_n to be $1 - \epsilon$ or $\epsilon/(2n)^2$ according as $n = 0$ or $n > 0$. Then (33) has positive coefficients, $f(w)$ is

regular for $|w| < 1$ but not at $w = 1$, and (19) is satisfied. It will be shown that if R is defined by

$$R = \int_0^{1-0} (f(r))^{-1} dr \tag{34}$$

(which implies that $R < \infty$), then the solution $w(z)$ of (6) and (2) must become singular at the point $z = R$. This will prove (β^*), since it will imply that $w(z)$ must become singular within the circle $|z| < 1/(1-\epsilon)$ (where the denominator can be chosen arbitrarily close to 1). In fact, since $f(r)$ is positive and increasing on the interval $0 \leqq r < 1$, it is clear from (34) that $R < 1/f(0)$, where $f(0) = c_0 = 1 - \epsilon$.

It follows from (2) and (6), where $f(0) \neq 0$, that (for small $|z|$) the function $w = w(z)$ is the (local) inverse of the function $z = z(w)$ defined (for small $|w|$) by

$$z = \int_0^{w} (f(t))^{-1} dt. \tag{35}$$

Consider the z-map of the interval for $0 \leqq w < 1$. Since the function $f(w) = f^*(w)$ is positive for $0 \leqq w < 1$ and has a positive and increasing derivative for $0 < w < 1$, it is clear from (35) and (34) that the interval $0 \leqq w < 1$ has the *schlicht* image $0 \leqq z < R$, and that, since (19) is satisfied, dz/dw stays between two positive bounds as $w \to 1$. Hence, if $w(z)$ did not become singular at $z = R$, then $z(w)$ would remain regular at $w = 1$. In view of (35), this is possible only if $1/f(w)$ remains regular at $w = 1$. Since $f(w)$ was chosen to be singular at $w = 1$, it follows that $f(w)$ has a pole at $w = 1$. But this contradicts (19).

Appendix.

In the assumptions of (i), the absolute value of the given function $f(z, w)$ is limited by 1 from *above*. If this limitation is made from *below*, the result is as follows:

(iii) *If $f(z, w)$ is regular, and satisfies the inequality $|f(z, w)| \geqq 1$, on the dicylinder $(|z| < 1, |w| < 1)$, then the solution $w(z)$ of* (1) *and* (6) *is regular, and satisfies the inequality $|w(z)| < 1$, on the circle $|z| < |f(0,0)|^{-2}$.*

In the proof of (iii), the following lemma (*) will be needed:

(*) *If* $w = w(z)$, *where* $|z| < 1$, *is any regular function satisfying the conditions* $w(0) = 0$, $w'(0) \neq 0$, *where* $w' = dw/dz$, *and the inequality* $|w'(z)| < 1$ *for* $|z| < 1$, *then the inverse function* $z = z(w)$ *is regular, and in absolute value less than* 1, *on the circle*

$$|w| < |w'(0)|^2. \tag{36}$$

As will be seen in a moment, (*) supplies not only (iii) but also the following fact:

(iii bis) *Under the assumptions and in the notations of* (iii), *the function* $w(z)$ *is schlicht in the circle* $|z| < |f(0,0)|^{-2}$.

Similarly, (i) can now be completed as follows:

(i bis) *Under the assumptions and in the notations of* (i), *the inverse,* $z = z(w)$, *of the function* $w(z)$ *is regular and schlicht in the circle* $|w| < |f(0,0)|^2$, *provided that* $f(0,0) \neq 0$.

In fact, (i bis) is clear from (i) and (*). It is also clear that both (iii) and (iii bis) follow from (*) and (i) if z and w are interchanged and, correspondingly, (6) and (2) are written as $dz/dw = F(w,z)$ and $z(0) = 0$, where $F(w,z) = 1/f(z,w)$ and $z = z(w)$. Thus it is sufficient to verify (*).

Remark. If $f(z,w)$ is of the particular form $f(w)$, as in (ii), then (i bis), where $|f(0,0)| < 1$ by assumption, can be completed as follows:

(ii bis) *Under the assumptions and in the notations of* (ii), *the function* $w(z)$ *is schlicht in the circle* $|z| < 1$, *provided that* $f(0) \neq 0$ (*if* $f(0) = 0$, *then* $w(z) \equiv 0$).

(ii bis) is a corollary of the sharper results on the local inverse of the integral (35) which are contained in a paper written in cooperation with Dr. Hartman (*Rend. Palermo,* ser. 2, vol. 3 (1954), pp. 286-292). But (ii bis) itself is trivial. For, on the one hand, the solution $w(z)$ of (6) and (2) is regular, and satisfies $|w(z)| < 1$, for $|z| < 1$ and, on the other hand, (35) defines z as a single-valued and regular function if w and the integration path joining w to 0 are confined to any simply-connected domain which is contained in the circle $|w| < 1$ and from which the zeros of $f(w)$

(if there are any in $|w| < 1$) have been excluded (by joining these zeros to a point of the circumference $|w| = 1$ by cuts).

Proof of (*). If the constants which in Satz X of Landau [3], p. 473, are denoted by M, R and a are chosen to be 1, 1 and $|w'(0)|$ respectively, it follows that the assertion of (*) is certainly true if the radius, $|w'(0)|^2$, of the circle (36), claimed in (*), is replaced by

$$1 + (|w'(0)|^{-2} - 1) \log (1 - |w'(0)|^2). \tag{37}$$

Hence (*) will be proved if it is ascertained that the value of (37), where $0 < |w'(0)| < 1$, exceeds $|w'(0)|^2$.

Clearly, $\log(1 - |w'(0)|^2) = -(|w'(0)|^2 + \frac{1}{2}|w'(0)|^4 + \cdots)$, where all the higher terms, those indicated by dots, have positive coefficients. Hence the value of (37) exceeds $1 + (1 - |w'(0)|^{-2})|w'(0)|^2$, which is $|w'(0)|^2$.

THE JOHNS HOPKINS UNIVERSITY.

REFERENCES.

PART I.

[1] A. Haar, "Über Eindeutigkeit und Analytizität der Lösungen partieller Differentialgleichungen," *Atti del Congresso Internazionale dei Matematici*, vol. 3 (1928), pp. 5-10.

[2] O. Hölder, "Abschätzungen in der Theorie der Differentialgleichungen," *Schwarz Festschrift* (1914), pp. 116-132.

[3] E. Landau, "Der Picard-Schottkysche Satz und die Blochsche Konstante," *Sitzungsberichte der Preussischen Akademie der Wissenschaften*, vol. 1926, pp. 467-474.

[4] E. Lindelöf, "Sur l'application des méthodes d'approximations successives à l'étude des intégrales réelles des équations différentielles ordinaires," *Journal de Mathématiques*, ser. 4, vol. 10 (1894), pp. 117-128.

[5] R. Lipschitz, *Differential- und Integralrechnung* (1880), pp. 506-511.

[6] M. Müller, "Über den Konvergenzbereich des Verfahrens der schrittweisen Näherungen bei gewöhnlichen Differentialgleichungen," *Mathematische Zeitschrift*, vol. 41 (1936), pp. 163-175.

[7] P. Painlevé, "Sur le calcul des intégrales d'un système différentiel par la méthode de Cauchy-Lipschitz," *Bulletin de la Societé Mathématique de France*, vol. 27 (1899), pp. 149-152.

[8] A. Wintner, "On the exact value of the bound for the regularity of solutions of ordinary differential equations," *American Journal of Mathematics*, vol. 57 (1935), pp. 539-540.

PART II.

[9] E. Landau, "Über die Blochsche Konstante und zwei verwandte Weltkonstanten," *Mathematische Zeitschrift*, vol. 30 (1929), pp. 608-634.

[10] ———, *Darstellung und Begründung einiger neuerer Ergebnisse der Funktionentheorie*, 2nd ed. (1929).

[11] E. Lindelöf, "Démonstration élémentaire de l'existence des fonctions implicites," *Bulletin des Sciences Mathématiques*, ser. 2, vol. 23 (1899), part I, pp. 68-75.

[12] H. Nakano, "Über den Konvergenzradius der Lösung einer Differentialgleichung $dy/dx = f(x, y)$," *Proceedings of the Imperial Academy of Japan*, vol. 8 (1932), pp. 29-31.

[13] I. Schur, "Über Potenzreihen, die im Innern des Einheitskreises beschränkt sind.," *Journal für die reine und angewandte Mathematik*, vol. 147 (1917), pp. 205-232 and vol. 148 (1918), pp. 122-145.

[14] A. Wintner, "On the bound of regularity of the solution of analytic differential equations of first order," *Quarterly Journal of Mathematics* (Oxford), ser. 2, vol. 5 (1954), pp. 145-149.

[15] ———, "Sur le calcul des limites de Cauchy dans la théorie des équations différentielles ordinaires," *Comptes Rendus*, vol. 242 (1956), pp. 1106-1107.

[16] ———, "On an absolute constant pertaining to Cauchy's 'principal moduli' in bounded power series," to appear in the *Mathematica Scandinavica* (1956).

ALGEBRAIC GROUPS OVER FINITE FIELDS.*

By Serge Lang.

1. Introduction. Let k be a finite field with q elements. Let G be an algebraic group defined over k. (For the foundations of the theory of algebraic groups and homogeneous spaces, see Weil [8], [9].) If x is a point of G, we denote by $x^{(q)}$ the point obtained by raising all coordinates of x to the q-th power, i. e. by applying to x the Frobenius automorphism of the universal domain leaving k fixed. The mapping $f(x) = x^{-1}x^{(q)}$ is a rational map of G into itself. It will be shown that it is in fact surjective, and that it gives a Galois, in general non abelian unramified covering of G over itself, the Galois group being that of the left rational translations. This sort of covering is of course impossible in characteristic 0.

More generally, it will be shown that G acts on itself as a homogeneous space, under the operation $F(x, y) = x \cdot y = x^{(q)}yx^{-1}$. Using this fact we shall show that every homogeneous space H of G defined over k has a rational point. If it is principal homogeneous, then it must be biregularly equivalent to G over k, and in case G is commutative, this means that the Weil group is trivial. We also use this result to give a new proof of a result due to Chatelet [4], that if a variety V/k becomes biregularly equivalent to projective space over the algebraic closure of k, then it is already so over k itself.

Finally we consider the class field theory for our covering defined by the map $z \to z^{-1}z^{(q)}$, get a partial non abelian reciprocity law, and prove that the Artin L-series are trivial. This is used to prove the following result: Let $\mathfrak{g}$ be a subgroup of the rational points of G over k. Let H be the homogeneous space of cosets of $G \bmod \mathfrak{g}$. Then G and H have the same number of rational points.

In case the group G is commutative, then one can get a complete reciprocity law, and one can use it to derive the class field theory over a variety V having a rational map into G by means of which the abelian coverings of G by commutative groups can be pulled back in a one-one manner.

* Received March 9, 1956.

This abelian class field theory is carried out in detail elsewhere, and in this paper, we have concentrated exclusively on the non abelian aspects of the covering $z^{-1}z^{(q)}$.

2. The map $f(z) = z^{-1}z^{(q)}$. The map $f(z)$ being as above we contend that if z is a generic point of G/k then $f(z)$ is also a generic point of G/k, and in fact the extension $k(z)$ over $k(f(z))$ is separable algebraic. Namely, putting $x = f(z)$ we have $k(z) = k(x, z) = k(x, z^{(q)})$. Our contention now follows from the following elementary and well known result of field theory:

PROPOSITION 1. *Let F be a field and E/F a finitely generated extension. We assume the characteristic p is $\neq 0$. If E/F is separable algebraic, then $E^{p^\mu}F = E$ for all powers p^μ, and conversely, if $E^{p^\mu}F = E$ for some power p^μ, then E/F is separably algebraic.*

(By E^{p^μ} we denote the field obtained by raising all elements of E to the p^μ-th power.)

We have trivially for z, w in G:

$$(1) \qquad (zw)^{(q)} = z^{(q)}w^{(q)}, \ (z^{-1})^{(q)} = (z^{(q)})^{-1}.$$

Furthermore our rational map satisfies the following formalism:

$$(2) \qquad f(zw) = f(z)^w f(w)$$

where y^w is defined to be $w^{-1}yw$.

The subgroup of G consisting of the elements rational over the field k_d (unique extension of k of degree d) will be denoted by G_d.

(3) An element a of G is in G_1 if and only if $f(a) = 0$. More generally, $f(z) = f(w)$ if and only if $z = aw$ for some a in G_1.

The first part of the statement is obvious. If $z = aw$, it is also obvious that $f(z) = f(w)$. Suppose $f(z) = f(w)$. Then $z^{-1}z^{(q)} = w^{-1}w^{(q)}$, whence $(zw^{-1}) = (zw^{-1})^{(q)}$. This means that zw^{-1} is rational over k, as desired.

We now see that if z is a generic point of G/k, then the extension $k(z)/k(f(z))$ is Galois, the distinct conjugates of z over $k(f(z))$ being az, with a rational over k.

The mapping $g(z) = z^{(q)}z^{-1}$ has analogous properties with respect to right translations, and we shall use them freely whenever needed.

PROPOSITION 2. *Let y be an arbitrary point of G and x a generic point over $k(y)$. Then $x^{(q)}yx^{-1}$ is a generic point of G over $k(y)$.*

Proof. Let $w = x^{(q)}yx^{-1}$. Put $K = k(y)$. Then $K(w, x) = K(w, x^{(q)})$. This implies that $K(x)$ is separable algebraic over $K(w)$, and that w has the same dimension as x over K. Hence it is a generic point of G/K.

Given two arbitrary points x and y of G, we denote by $x \cdot y$ the point $x^{(q)}yx^{-1}$. We obviously have for x, y, z arbitrary,

$$(xy) \cdot z = x \cdot (y \cdot z) \quad \text{and} \quad e \cdot x = x \tag{4}$$

To prove that G is a homogeneous space over itself with the above defined external law of composition, we need only prove the following statement:

THEOREM 1. *Given two points y and w in G, there exists a point x such that $x \cdot y = w$.*

Proof. In fact, using the associativity, it suffices to prove: Given y in G, there exists x such that $x \cdot y = e$. Let t be a generic point of G over $K = k(y)$. Then $t \cdot y$ is a generic point of G/K by Proposition 2. There is an isomorphism σ which is identity on K and maps $t^{(q)}t^{-1}$ on $t \cdot y = t^{(q)}yt^{-1}$. We can extend σ to the field $K(t)$. Let $u = t^{\sigma}$. Then

$$g(u) = g(t^{\sigma}) = g(t)^{\sigma} = t^{(q)}yt^{-1}.$$

If we put $x = u^{-1}t$, we have what we want.

COROLLARY. *The map $z \to z^{-1}z^{(q)}$ is surjective, i.e. given any y in G, there exists z such that $y = z^{-1}z^{(q)}$.*

Proof. According to the theorem, there exists z in G such that $z^{(q)}y^{-1}z^{-1} = e$. This element z does what is required.

From this corollary we see that we have indeed an unramified covering of G over itself. Given any point Q in G, there exists n points P such that $Q = P^{-1}P^{(q)}$, where n is the order of G_1. Given any one of them, all the others are simply the left rational translations of this one.

As another application of Theorem 1, we prove

THEOREM 2. *Let H/k be a homogeneous space over G. Then H has a rational point.*

Proof. We must show that there is some point u in H such that $u^{(q)} = u$. Let v be any point of H. Since H is defined over k, then $v^{(q)}$ is also in H. Since H is a homogeneous space, there exists an element x in G such that $xv^{(q)} = v$. By the corollary to Theorem 1, we can write $x = y^{-1}y^{(q)}$. From

this we see that $(yv)^{(q)} = (yv)$ and hence $u = yv$ is the element we are looking for.

We would like to point out here that Theorem 3 of [6] is a special case of the preceding result. Indeed, if a variety V/k becomes biregularly equivalent to an abelian variety over the algebraic closure of k, then V can be viewed as a principal homogeneous space over its Albanese variety A, which is known (by Chow's work) to be defined over k. (See Weil [9], Prop. 4.) However, because of the completeness of the group, we could give a direct proof, without using the Albanese variety. The proof can in fact be further simplified as follows:

We wish to prove that if a variety V over a finite field k becomes biregularly equivalent to a complete group variety over the algebraic closure of k, and hence over a finite extension k' of k, then V has a rational point over k. Over k' we can put a law of composition on V which makes it into a complete group variety (we don't even need to know it is abelian). There is a unit element e, rational over k', but not necessarily over k. Let z be a generic point of V over k. With respect to the composition law over k', we consider the point $f(z) = z^{-1}z^{(q)}$. It is a generic point of V over k' (by Proposition 1). Since V is complete, we can extend a specialization of $f(z)$ on e to a specialization of z to a point ξ on V, and then $\xi^{-1}\xi^{(q)} = e$. This point ξ satisfies $\xi = \xi^{(q)}$, and is therefore a rational point.

The statement made in our above mentioned paper that there is a rational place is a consequence of the following well known fact: *Let k be any field, and V/k any variety (say affine). Let Q be a simple point of V, rational over k. Let v be a generic point of V/k. Then there exists a place ϕ of $k(v)$ over k such that $\phi(v) = Q$ and such that ϕ is rational over k (i. e. k-valued).* One often says that Q is at the center of ϕ. Here of course, we have the additional property that the place can be chosen in such a way that the residue class field is canonically isomorphic to k itself: No irrationalities are needed in extending the specialization $v \to Q$ to a place of the function field. There exist many proofs of the above fact, and one of them runs along the following lines. The completion of the local ring of Q in $k(v)$ is isomorphic to a power series ring in r variables ($r = \dim V$) with coefficients in k, because Q is simple. There is therefore a canonical isomorphism of $k(v)$ in the quotient field of this power series ring. This quotient field itself can be embedded in the repeated power series field, which has obviously a k-valued place mapping all the variables on 0, one after the other. The restriction of this place to $k(v)$ is the one we are looking for.

We now return to the arbitrary algebraic group G over the finite field k.

Let $\sigma, \tau, \cdots$ range over the group of automorphisms of the extension k_d of k. This group (which is cyclic) operates on G_d in an obvious fashion. Referring to this operation, we show that the 1-cocycles split:

PROPOSITION 3. *Let $\{x_\sigma\}$ be a set of elements of G_d such that $x_\tau{}^\sigma x_\sigma = x_{\sigma\tau}$. Then there exists y in G_d such that $x_\sigma = y^\sigma y^{-1}$.*

Proof. We can change our indices from σ to integers mod d. According to the corollary of Theorem 2, we can write $x_1 = y^{(q)}y^{-1}$ for some y in G (we do not know yet whether it is in G_d). Then by the cocycle property, we get $x_i = y^{(q^i)}y^{-1}$. Finally, taking $i = d$ we must have

$$e = x_0 = x_d = y^{(q^d)}y^{-1}.$$

This shows that y is rational over k, because $y^{(q^d)} = y$.

As an application, we prove Chatelet's theorem:

THEOREM 3. *Let V be an abstract variety defined over k, which becomes biregularly equivalent to projective space S over the algebraic closure of k. Then V is biregularly equivalent to S over k.*

Proof. Let $T\colon V \to S$ be the correspondence, defined over some k_d. With Chatelet ([3], [4]) we take $x_\sigma = T^\sigma T^{-1}$, which is a birational biregular correspondence between S and itself. It is therefore projective, and is an element of the projective group G, rational over k_d. It satisfies the condition of Proposition 3, and if we let y be the projective transformation as in Proposition 3, we consider $T_1 = y^{-1}T$. Then T_1 is obviously fixed under every automorphism σ of k_d over k, and hence T_1 is defined over k.

For a proof that the only biregular correspondences of S with itself are projective, see Chow [5]. In that paper, other varieties are proved to have that property, and our theorem applies to them as well.

3. Class field theory. We shall now investigate the covering $f(x) = x^{-1}x^{(q)}$ from the point of view of class field theory. By a cycle, we shall always mean a cycle of dimension 0. Let $\mathfrak{p}$ be a prime rational cycle of G over k, and let Q be any point in it. Let P be any point such that $f(P) = Q$. All other points are of type aP, where a is rational over k. If $\mathfrak{p}$ is or degree d, then $f(P^{(q^d)}) = f(P)^{(q^d)} = Q$. Hence $P^{(q^d)}$ also lies in the inverse image of Q under f, and hence there exists a rational point a in G_1 such that $aP = P^{(q^d)}$.

We define $\pi_d(Q)$ to be the product $QQ^{(q)}\cdots Q^{(q^{d-1})}$. Then we have

$$aP = P^{(q^d)} = P\pi_d(Q). \tag{6}$$

It is clear that the point a in G_1 is completely well defined by the prime $\mathfrak{p}$, up to conjugacy in G_1. Furthermore we have

PROPOSITION 4. *Given two points Q_1 and Q_2 in G_1, let a_1 and a_2 in G_1 be any points determined by the condition*

$$a_1P_1 = P_1{}^{(q)} = P_1Q_1, \qquad a_2P_2 = P_2{}^{(q)} = P_2Q_2.$$

Then a_1 is conjugate to a_2 in G_1 if and only if Q_1 is conjugate to Q_2 in G_1.

The proof is trivial and formal, and we leave it to the reader.

We have a mapping from the primes to the conjugacy classes of G_1 defined by (6), and if the prime is of degree 1, then we obtain a 1-1 mapping of the conjugacy classes of G_1 (viewed as a set of primes of degree one) onto the conjugacy classes of G_1 (viewed as Galois group of left translations). The period of the point a is clearly equal to the period of $\pi_d(Q)$, and thus we can tell the period of the Frobenius class associated with a prime. Furthermore, we see that $\mathfrak{p}$ splits completely if and only if $\pi_d(Q) = e$.

I have not been able to determine whether the conjugacy class of a is always equal to that of Q (when Q is in G_1) and more generally to determine if a and $\pi_d(Q)$ are conjugate in G_d. In order to obtain the complete decomposition laws, we would need to know that any rational point b in G_1, conjugate to $\pi_d(Q)$ in G_d is conjugate to a in G_1. This would imply that the Frobenius class associated with $\mathfrak{p}$ in G_1 can be determined rationally. (This is the case for the full linear group.)

We now consider the L-series.

Let z be a generic points of G/k. Let

$$f_n(z) = z^{-1}z^{(q^n)} \text{ and } \pi_n(z) = zz^{(q)}\cdots z^{(q^{n-1})}$$

Then $\pi_n(f_1(z)) = f_n(z)$.

Let $E = k(z)$, $F = k(f_1(z))$, and $K = k(f_n(z))$. We have inclusions $E \supset F \supset K$, and E/F is Galois (it is the extension discussed previously). Let E_n, F_n, and K_n be the constant field extensions by k_n. Then E_n/K_n becomes Galois, with group G_n. The group G is a model of each one of our function fields, but of course in a different way each time.

Referring to these models, we shall prove that the L-series are trivial for a character not containing the identity.

Let χ be a character of G_1. Let $\mathfrak{p}$ be a prime cycle of G/k. Let $T_{\mathfrak{p}}$ be any one of the translations of G_1 in the Frobenius class associated with $\mathfrak{p}$ in G_1. Then the L-series are defined by

$$t\frac{d}{dt}\operatorname{Log} L(t,\chi,E/F) = \sum \chi(T_{\mathfrak{p}}{}^{\mu})\deg(\mathfrak{p})\, t^{\mu \deg(\mathfrak{p})}$$

the sum being taken over all primes and all $\mu \geqq 1$.

If we look at the coefficient of t^n we see that we must prove that

$$\sum_{\deg(\mathfrak{p})\mid n} \chi(T_{\mathfrak{p}}{}^{n/\deg(\mathfrak{p})})\deg(\mathfrak{p}) = 0.$$

This sum can be rewritten in terms of points in G_n as follows: To each point Q in G_n, we can associate a translation $T_Q{}^{(n)}$ (well defined up to conjugacy) in G_1, such that for any P in $f_1{}^{-1}(Q)$ we have $T_Q{}^{(n)}(P) = P^{(q^n)}$. It is then clear that the above sum is equal to

$$\sum_{Q \in G_n} \chi(T_Q{}^{(n)}). \tag{7}$$

For $n = 1$ we know by Proposition 4 that each class will have a representative $T_Q{}^{(1)}$ for some Q in G_1, and that this representative will occur in our sum as many times as there are element in that class. Consequently the value of our sum is the same as the value of the character taken over the sum of the conjugacy classes in the group ring of G_1. If χ does not contain the identity character, then this sum must be 0.

For arbitrary n, we note that the coefficient of t^n in our L-series $L(t,\chi,E/F)$ is by definition and formula (7) the same as the coefficient of t in $L(t,\chi,E_n/F_n)$. We have thus reduced the computation of the n-th term to the computation of the first term of an L-series over a bigger constant field. By one of the main theorems on L-series, we know that

$$L(t,\chi,E_n/F_n) = L(t,\chi^*,E_n/K_n)$$

where χ^* is the induced character. If χ does not contain the identity, then neither does χ^*. The extension E_n/K_n is now of a type analogous to that considered above for $n = 1$ (i.e. belonging to a rational map f_n). Hence the first coefficient of the L-series is equal to zero, as desired.

Let $\lambda\colon G \to H$ be a homomorphism of an algebraic group G onto an algebraic group H, defined over k, and with finite kernel. Let z be a generic point of G/k. As usual, $f_G(z) = z^{-1}z^{(q)}$. We also have an f-mapping on H, denoted by f_H. Then obviously $f_H\lambda = \lambda f_G$. Taking the degree of both sides, we see that the degree of f_H must equal that of f_G, i.e. that *G and H have*

the same number of rational points (hence the same zeta function). If kernel of λ is contained in G_1, and λ is separable, then G is Galois over H, and this again suggests that the L-series for the covering are trivial. This is indeed the case, and can be proved as follows: Put $y = \lambda(z)$. Then $E = k(z) \supset k(y) = M \supset k(z^{-1}z^{(q)}) = F$. If χ is a character for the Galois group of E/M, and does not contain the identity, then the induced character χ^* to the Galois group of E/F does not contain the identity either, and by what has been proved before, the L-series belonging to it must be trivial.

More generally, if the identity for E/M occurs with some multiplicity in χ, then the identity for E/F occurs with the same multiplicity in χ^*. (See for instance Brauer-Tate [2], where we put $\Theta =$ identity in formula (5).) From this remark we can deduce the following result.

Let H be the homogeneous space obtained from G by the cosets of a subgroup of G_1. Then we have a rational map λ: $G \to H$ (not necessarily a homomorphism), and the same type of field inclusion as before: $k(z) \supset k(y) \supset k(x)$. The zeta function of H, denoted by $Z_H(t)$, can be written

$$Z_H(t) = L(t, 1, E/M) = L(t, 1^*, E/F)$$

where 1 stands for the identity character on the Galois group of E/M. The identity for E/F occurs exactly once in 1^*, and hence the above L-series is equal to $Z_G(t) \cdot L(t, \chi, E/F)$, where χ is some character for E/F, which does not contain the identity, and $Z_G(t)$ is the zeta function of G. This shows that the zeta function of G and H coincide, and hence that G and H have the same number of rational points.

Knowing that the L-series are trivial, we can of course apply the formal argument given by Artin (Satz 4 of [1]) to get the density of primes in a given arithmetic progression. Let χ_i be the simple characters of G_1. Then we know that

$$\sum_{Q \in G_n} \chi_i(T_Q{}^{(n)}) = \begin{cases} 0 & i \neq 1 \\ q^{nr} + O(q^{n(r-1/2)}) & i = 1 \end{cases} \tag{8}$$

because for $i = 1$, we deal with the zeta function and can use the results of [7]. Let C_j be a fixed class in G_1 and T any element of C_j. Let h be the order of G_1 and h_j the number of elements in C_j. Multiplying (8) by $\chi_i(T^{-1})$ and summing, we use the orthogonality relations (see formula (3) of [1]) to get

$$\frac{h}{h_j} N(n, C_j) = q^{nr} + O(q^{n(r-1/2)}),$$

where $N(n, C_j)$ is the number of points in G_n having their $T_Q{}^{(n)}$ in the given class C_j. One sees trivially that to get an estimate for the number of primes, one has to divide by n.

COLUMBIA UNIVERSITY.

REFERENCES.

[1] E. Artin, "Über eine neue Art von L-Reihen," *Abhandlungen aus dem Mathematischen Seminar der Hamburgischen Universität*, vol. 3 (1923), pp. 89-108.

[2] R. Brauer and J. Tate, "On the characters of finite groups," *Annals of Mathematics*, vol. 62 (1955), pp. 1-7.

[3] F. Chatelet, "Variations sur un thème de Poincaré," *Annales de l'Ecole normale superieure*, vol. 61 (1944), pp. 249-300.

[4] ———, "Les courbes de genre 1 dans un champ de Galois," *Comptes rendus des seances de l'Academie des Sciences*, vol. 224 (1947), pp. 1616-1618.

[5] W. L. Chow, "On the geometry of algebraic homogeneous spaces," *Annals of Mathematics*, vol. 50 (1949), pp. 32-67.

[6] S. Lang, "Abelian varieties over finite fields," *Proceedings of the National Academy of Sciences*, vol. 41 (1955), pp. 174-176.

[7] ——— and A. Weil, "Number of points of varieties in finite fields," *American Journal of Mathematics*, vol. 76 (1954), pp. 819-827.

[8] A. Weil, "Algebraic groups of transformations," *American Journal of Mathematics*, vol. 77 (1955), pp. 355-391.

[9] ———, "On algebraic groups and homogeneous spaces," *American Journal of Mathematics*, vol. 77 (July 1955), pp. 493-512.

REPRESENTATIONS OF SEMISIMPLE LIE GROUPS VI.*

Integrable and Square-Integrable Representations.

By HARISH-CHANDRA.

1. Introduction. Let G be a connected semisimple Lie group and let Z denote its center. Then if π is an irreducible unitary representation of G on a Hilbert space $\mathfrak{H}$ we can find a unitary character η_π of Z such that $\pi(z) = \eta_\pi(z)\pi(1)$ for $z \varepsilon Z$. Let $x \to x^*$ denote the natural mapping of G on $G^* = G/Z$. If $\phi, \psi \varepsilon \mathfrak{H}$, it is clear that $(\phi, \pi(x)\psi)\,(x \varepsilon G)$ depends only on x^*. Let dx^* denote the Haar measure on G^*. We shall say that π is square-integrable if there exists an element $\psi_0 \neq 0$ in such that $\int_{G^*} |(\psi_0, \pi(x)\psi_0)|^2\, dx^* < \infty$. Similarly π is said to be integrable if $\int_{G^*} |(\psi_0, \pi(x)\psi_0)|\, dx^* < \infty$ for some $\psi_0 \neq 0$ in $\mathfrak{H}$. In this paper we intend to study in detail some examples of such representations.

Let π be a square-integrable representation of G. Then, as shown by Godement [4(b)], the Schur orthogonality relations hold for π and therefore there exists a positive constant d_π such that

$$\int_{G^*} |(\phi, \pi(x)\psi)|^2\, dx^* = d_\pi^{-1} |\phi|^2 |\psi|^2$$

for all $\phi, \psi \varepsilon \mathfrak{H}$. Naturally d_π depends on the normalization of the measure dx^* but once this has been fixed, d_π can be considered as a function of π. In analogy with the case of compact groups we call d_π the formal degree of π. If Z is finite and ω is the equivalence class of π, d_π is also equal to the mass of ω with respect to the Plancherel measure (see Section 5) just as in the compact case. Moreover for compact semisimple groups Weyl [11(a)] has given a formula for the degree of an irreducible representation in terms of its "highest weight." We shall se that substantially the same formula holds for the formal degree d_π under suitable conditions. This is the principal result of this paper. It can be verified immediately in the case of the 2×2 real unimodular group by looking at the results obtained

* Received September 27, 1955.

by Bargmann [1, p. 634] by direct computation. This very simple case is, in a sense, fundamental and much of our argument will depend on the properties of this three-dimensional group.

This paper is divided into two parts. In Part I we obtain the Schur orthogonality relations and establish the connection between the formal degree and the Plancherel measure. We also obtain a formula for the character of a square-integrable representation which is quite similar to the corresponding formula in the compact case. Although the Schur orthogonality relations are now new, Godement's proof of them [4(b)] does not cover the case when Z is infinite. (However it could perhaps be modified to include this case as well.) At any rate our method is quite different and it seems worthwhile to present it briefly even at the risk of some overlap with earlier work especially since the infinite case mentioned above is particularly important for us.

Part II is devoted to the proof of the analogue of Weyl's formula. This proof depends on a detailed comparison at each step between the compact and the non-compact cases and the entire argument is based on Lemma 22. In order to make this comparison we need some considerable algebraic preparation which consists of an intensive study of the root-structure of certain types of semisimple Lie algebras. As an incidental outcome of this study, we get in Section 7 a new proof of a theorem of E. Cartan [2(c), p. 145)] on the boundedness of certain complex domains. This proof, unlike that of Cartan, does not depend on the classification of simple groups.

The last two sections contain the proof of Lemma 22. Here I follow closely a method due to Weyl [11(a)] and Cartan [2(a)] and although the proof is somewhat long no new ideas are involved.

The results of this paper had been announced in a short note [5(d)].

Part I.

2. Preliminary lemmas. Let G be a connected semisimple Lie group and let $\mathfrak{g}_0$ denote its Lie algebra over the field R of real numbers. Define $\mathfrak{k}_0$ as in [5(b)] and let $\mathfrak{c}_0$ be the center and $\mathfrak{k}_0' = [\mathfrak{k}_0, \mathfrak{k}_0]$ the derived algebra of $\mathfrak{k}_0$. Let K, K' and D denote the analytic subgroups of G corresponding to $\mathfrak{k}_0$, $\mathfrak{k}_0'$ and $\mathfrak{c}_0$ respectively. We consider the space $C_c(G)$ of all (complex-valued) continuous functions on G which vanish outside a compact set. For any two functions $f, g \in C_c(G)$ we define their convolution $f * g$ by

$$(f * g)(x) = \int_{G^*} f(y) g(y^{-1}x) dy \qquad (x \in G)$$

where dy is the Haar measure on G. Under this operation $C_c(G)$ becomes an associative algebra. Let Ω denote the set of all equivalence classes of finite-dimensional irreducible representations of K and let $\xi_{\mathfrak{D}}$ be the character (on K) of any class $\mathfrak{D}\in\Omega$. Choose a base $\Gamma_1,\cdots,\Gamma_r$ for $\mathfrak{c}_0$ over R such that $\exp\Gamma_i\in D\cap Z$, $1\leqq i\leqq r$. (Z is the center of G.) This is possible since $D/D\cap Z$ is compact (see Mostow [9]). Let $\mathfrak{c}_1$ be the subset of $\mathfrak{c}_0$ consisting of all elements of the form $t_1\Gamma_1+\cdots+t_r\Gamma_r$ ($t_i\in R$) with $|t_i|\leqq\frac{1}{2}$, $1\leqq i\leqq r$. Then $K_0=K'(\exp\mathfrak{c}_1)$ is a compact subset of K. For any $f\in C_c(G)$ and $\mathfrak{D}\in\Omega$ we define two functions ${}_{\mathfrak{D}}f$ and $f_{\mathfrak{D}}$ in $C_c(G)$ as follows:

$$ {}_{\mathfrak{D}}f(x)=d(\mathfrak{D})\int_{K_0}\xi_{\mathfrak{D}}(u)f(ux)\,du,\qquad f_{\mathfrak{D}}(x)=d(\mathfrak{D})\int_{K_0}f(xu)\xi_{\mathfrak{D}}(u)\,du $$

($x\in G$). Here du is the Haar measure on K normalized in such a way that $\int_{K_0}du=1$. Also $d(\mathfrak{D})$ is the degree of any representation in $\mathfrak{D}$. Let $L(\mathfrak{D})$ ($\mathfrak{D}\in\Omega$) denote the subspace of $C_c(G)$ consisting of all functions of the form ${}_{\mathfrak{D}}f$ ($f\in C_c(G)$). Since ${}_{\mathfrak{D}}(f*g)=({}_{\mathfrak{D}}f)*g$ ($f,g\in C_c(G)$ it follows that $L(\mathfrak{D})$ is a right ideal in $C_c(G)$.

Let π be a quasi-simple irreducible representation (see [5(b)] of G on a Banach space $\mathfrak{H}$. For any $\mathfrak{D}\in\Omega$, let $\mathfrak{H}_{\mathfrak{D}}$ denote the subspace of $\mathfrak{H}$ consisting of all those elements which transform under $\pi(K)$ according to $\mathfrak{D}$. Let Ω_π be the set of all $\mathfrak{D}\in\Omega$ such that $\mathfrak{H}_{\mathfrak{D}}\neq 0$. Then $L_\pi=\sum_{\mathfrak{D}\in\Omega_\pi}L(\mathfrak{D})$ is a subalgebra of $C_c(G)$. Since π is quasi-simple there exists a character η_π of Z such that $\pi(z)=\eta_\pi(z)\pi(1)$ ($z\in Z$). We shall call η_π *the central character of* π. For any $f\in C_c(G)$ let $\pi(f)$ denote the operator $\int_G f(x)\pi(x)\,dx$. Then $f\to\pi(f)$ defines a representation of $C_c(G)$ on $\mathfrak{H}$.

LEMMA 1. *The space* $\mathfrak{H}_0=\sum_{\mathfrak{D}\in\Omega}\mathfrak{H}_{\mathfrak{D}}$ *is invariant and (algebraically) irreducible under* $\pi(L_\pi)$.

Let $E_{\mathfrak{D}}$ ($\mathfrak{D}\in\Omega$) denote the canonical projection [5(b), p. 225] of $\mathfrak{H}$ on $\mathfrak{H}_{\mathfrak{D}}$. Then if $\mathfrak{D}\in\Omega_\pi$, one proves easily (see [5(c), p. 249]) that

$$ E_{\mathfrak{D}}=d(\mathfrak{D})\int_{K_0}\xi_{\mathfrak{D}}(u^{-1})\pi(u)\,du $$

and therefore

$$ E_{\mathfrak{D}}\pi(f)=\pi({}_{\mathfrak{D}}f)\qquad (f\in C_c(G)). $$

On the other hand if $\mathfrak{D}\notin\Omega_\pi$, $E_{\mathfrak{D}}=0$. Hence it is clear that if $f\in L_\pi$, $\pi(f)$ maps $\mathfrak{H}$ into $\mathfrak{H}_0$ Let ψ_0 be any nonzero element in $\mathfrak{H}_0$. In order to prove

the irreducibility of $\mathfrak{H}_0$ under $\pi(L_\pi)$ it would be enough to show that $\mathfrak{H}_{\mathfrak{D}} \subset \pi(L_\pi)\psi_0$ for all $\mathfrak{D} \varepsilon \Omega_\pi$. Suppose then that this is false for some $\mathfrak{D}$. Put $U = \mathfrak{H}_{\mathfrak{D}} \cap \pi(L_\pi)\psi_0$. Then $U \neq \mathfrak{H}_{\mathfrak{D}}$ and since $\dim \mathfrak{H}_{\mathfrak{D}}$ is finite [5(b)], there exists a linear function $\alpha \neq 0$ on $\mathfrak{H}_{\mathfrak{D}}$ which vanishes identically on U. Extend α to a continuous linear function on $\mathfrak{H}$ by setting $\alpha(\psi) = \alpha(E_{\mathfrak{D}}\psi)$ $(\psi \varepsilon \mathfrak{H})$. Since $\psi_0 \neq 0$ and π is an irreducible representation of G, it is obvious that the continuous function $\alpha(\pi(x)\psi_0)$ $(x \varepsilon G)$ cannot be everywhere zero on G. Hence we can choose $f \varepsilon C_c(G)$ such that

$$\int f(x)\alpha(\pi(x)\psi_0)\,dx \neq 0.$$

Then

$$\alpha(\pi({}_{\mathfrak{D}}f)\psi_0) = \alpha(E_{\mathfrak{D}}\pi(f)\psi_0) = \alpha(\pi(f)\psi_0) = \int f(x)\alpha(\pi(x)\psi_0)\,dx \neq 0.$$

But since ${}_{\mathfrak{D}}f \varepsilon L(D)$, $\pi({}_{\mathfrak{D}}f)\psi_0 \varepsilon U$ and therefore $\alpha(\pi({}_{\mathfrak{D}}f)\psi_0) = 0$. This contradiction proves the lemma.

COROLLARY. *For any $\mathfrak{D} \varepsilon \Omega_\pi$, $\mathfrak{H}_{\mathfrak{D}}$ is invariant and irreducible under $\pi(L(\mathfrak{D}))$ and the corresponding representation of $L(\mathfrak{D})$ on $\mathfrak{H}_{\mathfrak{D}}$ determines π completely up to infinitesimal equivalence* [5(b), p. 230].

Since $E_{\mathfrak{D}}\pi(f) = \pi({}_{\mathfrak{D}}f)$ $(f \varepsilon C_c(G))$ it is obvious that $\mathfrak{H}_{\mathfrak{D}}$ is invariant under $\pi(L(\mathfrak{D}))$. Let ψ_0 and ψ be two elements in $\mathfrak{H}_{\mathfrak{D}}$ and suppose $\psi_0 \neq 0$. Then it follows from the above lemma that $\psi = \pi(f)\psi_0$ for some $f \varepsilon L_\pi$. But since $\psi \varepsilon \mathfrak{H}_{\mathfrak{D}}$,

$$\psi = E_{\mathfrak{D}}\psi = \pi({}_{\mathfrak{D}}f)\psi_0.$$

However ${}_{\mathfrak{D}}f \varepsilon L(\mathfrak{D})$ and so the irreducibility is proved.

Let $\phi(x)$ $(x \varepsilon G)$ denote the trace of the restriction of $E_{\mathfrak{D}}\pi(x)E_{\mathfrak{D}}$ on $\mathfrak{H}_{\mathfrak{D}}$. Then if $f \varepsilon C_c(G)$, $\int f(x)\phi(x)\,dx$ is the trace of the restriction of $E_{\mathfrak{D}}\pi(f)E_{\mathfrak{D}}$ on $\mathfrak{H}_{\mathfrak{D}}$. But since $E_{\mathfrak{D}}\pi(f) = \pi({}_{\mathfrak{D}}f)$, it follows that

$$\int f(x)\phi(x)\,dx = \int {}_{\mathfrak{D}}f(x)\phi(x)\,dx.$$

Hence if ν denotes the representation of $L(\mathfrak{D})$ on $\mathfrak{H}_{\mathfrak{D}}$

$$\int f(x)\phi(x)\,dx = \mathrm{Sp}(\nu({}_{\mathfrak{D}}f)) \qquad (f \varepsilon C_c(G)).$$

ϕ being a continuous function, it is now obvious that the knowledge of the trace of ν determines it completely. From this our assertion follows (see [5(c), p. 235]).

LEMMA 2. *Let $\psi \neq 0$ be an element in $\mathfrak{H}$. Then $E_{\mathfrak{D}}\psi \neq 0$ for some $\mathfrak{D} \varepsilon \Omega_\pi$.*

Let $C_c^\infty(G)$ be the set of all functions $f \varepsilon C_c(G)$ which are everywhere indefinitely differentiable. Choose a neighborhood V of 1 in G such that $|\pi(x)\psi - \psi| \leqq \frac{1}{2}|\psi|$ for $x \varepsilon V$. Since K_0 is compact, we can find another such neighborhood V' with the property that $uV'u^{-1} \subset V$ for $u \varepsilon K_0$. Select a real-valued function $g \varepsilon C_c^\infty(G)$ such that $g \geqq 0$, $\int g(x)\,dx = 1$ and $g = 0$ outside V'. Then if

$$f(x) = \int_{K_0} g(uxu^{-1})\,du \qquad (x \varepsilon G)$$

it is obvious that $f \varepsilon C_c^\infty(G)$, $f \geqq 0$, $\int f(x)\,dx = 1$ and $f = 0$ outside V. Moreover since $K = K_0 Z$, one proves easily that $f(ux) = f(xu)$ $(u \varepsilon K, x \varepsilon G)$. Therefore $E_{\mathfrak{D}}\pi(f) = \pi(f)E_{\mathfrak{D}}$ and

$$|\pi(f)\psi - \psi| = |\int f(x)(\pi(x)\psi - \psi)dx| \leqq \int f(x)|\pi(x)\psi - \psi|\,dx \leqq \tfrac{1}{2}|\psi|.$$

Hence $\pi(f)\psi \neq 0$ and so it follows from Lemma 3 of [5(c)] that $\pi(f)E_{\mathfrak{D}}\psi = E_{\mathfrak{D}}\pi(f)\psi \neq 0$ for some $\mathfrak{D} \varepsilon \Omega_\pi$. This proves that $E_{\mathfrak{D}}\psi \neq 0$.

COROLLARY. *Let π' be another quasi-simple irreducible representation of G on a Banach space $\mathfrak{H}'$ and let $\psi' \neq 0$ be an element in $\mathfrak{H}'$. Suppose $\pi'(f)\psi' = 0$ whenever $\pi(f) = 0$ $(f \varepsilon C_c(G))$. Then π and π' are infinitesimally equivalent.*

Let η be the central character of π. For any $z \varepsilon Z$ and $f \varepsilon C_c(G)$, define a function ${}_zf \varepsilon C_c(G)$ by ${}_zf(x) = f(z^{-1}x)$ $(x \varepsilon G)$. Then $\pi({}_zf) = \eta(z)\pi(f)$ and therefore

$$\pi'(z)\pi'(f)\psi' = \pi'({}_zf)\psi' = \eta(z)\pi'(f)\psi' \qquad (z \varepsilon Z, f \varepsilon C_c(G)).$$

Since elements of the form $\pi(f)\psi'$ $(f \varepsilon C_c(G))$ are dense in $\mathfrak{H}'$, we conclude that η is also the central character of π'.

Now $Z \subset K$ and therefore, by Schur's lemma, there exists, for each $\mathfrak{D} \varepsilon \Omega$, a character $\eta_{\mathfrak{D}}$ of Z such that $\sigma(z) = \eta_{\mathfrak{D}}(z)\sigma(1)$ $(z \varepsilon Z)$ for any representation σ in $\mathfrak{D}$. Let Ω_0 be the set of all those $\mathfrak{D}$ for which $\eta_{\mathfrak{D}} = \eta$. Then it is obvious that $\Omega_\pi \cup \Omega_{\pi'} \subset \Omega_0$ and if $E_{\mathfrak{D}}'$ is the canonical projection of $\mathfrak{H}'$ and $\mathfrak{H}_{\mathfrak{D}}'$,

$$E_{\mathfrak{D}} = d(\mathfrak{D})\int_{K_0} \xi_{\mathfrak{D}}(u^{-1})\pi(u)\,du, \qquad E_{\mathfrak{D}}' = d(\mathfrak{D})\int_{K_0} \xi_{\mathfrak{D}}(u^{-1})\pi'(u)\,du$$

for all $\mathfrak{D} \varepsilon \Omega_0$ (see [5(c), p. 249]). Now, by the above lemma, we can choose $\mathfrak{D}_0 \varepsilon \Omega_{\pi'}$ such that $E_{\mathfrak{D}_0}'\psi' \neq 0$. Then $\pi'(f) E_{\mathfrak{D}_0}'\psi' \neq 0$ for some $f \varepsilon C_c(G)$. This implies that $\pi'(f_{\mathfrak{D}_0})\psi' \neq 0$ and therefore $\pi(f)E_{\mathfrak{D}_0} = \pi(f_{\mathfrak{D}_0}) \neq 0$. Hence $E_{\mathfrak{D}_0} \neq 0$ and so $\mathfrak{D}_0 \varepsilon \Omega_\pi \cap \Omega_{\pi'}$. Let ν and ν' be the corresponding representations of $L(\mathfrak{D}_0)$ on $\mathfrak{H}_{\mathfrak{D}_0}$ and $\mathfrak{H}_{\mathfrak{D}_0}'$ respectively. In view of the corollary to Lemma 1, it is sufficient to prove that ν and ν' are equivalent. But since they are both irreducible and finite-dimensional it would be enough to show that they have the same kernel. Let $\mathfrak{M}$ be the vernel of ν in $L(\mathfrak{D}_0)$. Then it is a maximal two-sided ideal in $L(\mathfrak{D}_0)$. If $g \varepsilon \mathfrak{M}$, $\pi(g_{\mathfrak{D}_0}) = \pi(g)E_{\mathfrak{D}_0} = 0$ and therefore $\pi'(g)E_{\mathfrak{D}_0}'\psi' = \pi'(g_{\mathfrak{D}_0})\psi' = 0$. This shows that $\nu'(\mathfrak{M})E_{\mathfrak{D}_0}'\psi' = 0$. Since $E_{\mathfrak{D}_0}'\psi' \neq 0$ and ν' is irreducible, it follows that the kernel of ν' must contain $\mathfrak{M}$ and therefore coincide with it. This completes the proof.

Let Z_0 be a fixed subgroup of Z such that Z/Z_0 is finite and let $x \to x^*$ denote the natural mapping of G and $G^* = G/Z_0$. Suppose $\alpha_1, \cdots, \alpha_r$ is a finite set of continuous linear functions on $\mathfrak{H}$ and $\psi_1, \cdots, \psi_r$ are certain given elements in $\mathfrak{H}$. Put

$$f(x) = \alpha_1(\pi(x)\psi_1) + \cdots + \alpha_r(\pi(x)\psi_r) \qquad (x \varepsilon G).$$

Then if the central character of π is unitary, it is obvious that $|f(x)|$ depends only on on x^*. Let dx^* denote the Haar measure on G^*.

LEMMA 3. *Assume that the function*

$$f(x) = \alpha_1(\pi(x)\psi_1) + \cdots + \alpha_r(\pi(x)\psi_r) \qquad (x \varepsilon G)$$

is not identically zero on G. Then if the central character of π is unitary and $\int_{G^*} |f(x)|^2 dx^* < \infty$, π is infinitesimally equivalent to an irreducible unitary repersentation of G on a Hilbert space.

Let η_π denote the central character of π. Then $f(xz) = \eta_\pi(z)f(x)$ $(x \varepsilon G, z \varepsilon Z)$ and therefore it follows easily from the Peter-Weyl theorem for the compact group $K/D \cap Z$ that

$$\int_{K_0} |f(xu)|^2 du = \sum_{\mathfrak{D} \varepsilon \Omega_\pi} \int_{K_0} |f_{\mathfrak{D}}(xu)|^2 du.$$

Since $f \neq 0$, we can choose $\mathfrak{D}_0 \varepsilon \Omega_\pi$ such that $f_{\mathfrak{D}_0} \neq 0$. Then it is clear that

$$f_{\mathfrak{D}_0}(x) = \sum_{i=1}^{r} \alpha_i(\pi(x)E_{\mathfrak{D}_0}\psi_i)$$

and since

$$\int_{K_0} |f_{\mathfrak{D}_0}(xu)|^2 du \leqq \int_{K_0} |f(xu)|^2 du,$$

it follows that

$$\int |f_{\mathfrak{D}_0}(x)|^2\,dx^* \leqq \int |f(x)|^2\,dx^* < \infty.$$

Hence if we replace ψ_i by $E_{\mathfrak{D}_0}\psi_i$ and f by $f_{\mathfrak{D}_0}$, our problem is reduced to the case when the given elements of $\mathfrak{H}$ all lie in $\mathfrak{H}_{\mathfrak{D}_0}$. Moreover it is obvious that without loss of generality we may assume that $\psi_1, \cdots, \psi_r$ is a base for $\mathfrak{H}_{\mathfrak{D}_0}$ over the field C of complex numbers. Since $L(\mathfrak{D}_0)$ is irreducible under $\pi(L(\mathfrak{D}_0))$ (Corollary to Lemma 1), it follows from Burnside's Theorem that

$$\pi(g_j)\psi_k = \delta_{jk}\psi_k \qquad 1 \leqq j, k \leqq r$$

for suitable elements $g_j \varepsilon L(\mathfrak{D}_0)$. ($\delta_{jk}$ is the Kronecker symbol.) Since $f \neq 0$ we may assume that $f_1(x) = \alpha_1(\pi(x)\psi_1)$ is not identically zero. Then

$$f_1(x) = \sum_{i=1}^{r} \alpha_i(\pi(x)\pi(g_1)\psi_i) = \int_G f(xy)g_1(y)\,dy$$

and therefore

$$|f_1(x)|^2 \leqq \int_\omega |f(xy)|^2\,dy \int_G |g_1(y)|^2\,dy$$

where ω is some compact subset of G outside which g_1 is zero. This shows that $\int |f_1(x)|^2\,dx^* < \infty$ and since $f_1 \neq 0$, our problem is now reduced to the case when

$$f(x) = \alpha(\pi(x)\psi).$$

Here α is a continuous linear function on $\mathfrak{H}$ and $\psi \varepsilon \mathfrak{H}_{\mathfrak{D}_0}$.

Let U' be the Hilbert space consisting of all measurable functions g on G such that (1) $g(xz) = g(x)\eta_\pi(z)$ $(x \varepsilon G, z \varepsilon Z)$ and (2) $\int_{G^*} |g(x)|^2 dx^* < \infty$. We define a representation σ' of G on U' by $(\sigma'(y)g)(x) = g(xy)$ $(x, y \varepsilon G,$ $g \varepsilon U')$. It is obvious that f lies in U'. Let U be the smallest closed subspace of U' containing f which is invariant under $\sigma'(G)$. We denote by σ the representation of G defined on U under σ'. It is clear that σ is unitary. We shall now show that π is infinitesimally equivalent to σ.

For any $\mathfrak{D} \varepsilon \Omega$ let $F_{\mathfrak{D}}$ denote the corresponding canonical projection in U. Also, we denote the operator $\int g(x)\sigma(x)\,dx$ $(g \varepsilon C_c(G))$ by $\sigma(g)$. Since $f \varepsilon U_{\mathfrak{D}_0}$, it follows that $F_{\mathfrak{D}_0} \neq 0$ and therefore $F_{\mathfrak{D}_0}\sigma(g) = \sigma({}_{\mathfrak{D}_0}g)$. Moreover it is obvious that elements of the form $\sigma(g)f$ $(g \varepsilon C_c(G))$ are dense in U. Therefore $\sigma(L(\mathfrak{D}_0))f$ is dense in $U_{\mathfrak{D}_0} = F_{\mathfrak{D}_0}U$. Now suppose $\pi(g)\psi = 0$ $(g \varepsilon C_c(G))$. Then

$$0=\alpha(\pi(x)\pi(g)\psi)=\int \alpha(\pi(xy)\psi)g(y)dy=\int f(xy)g(y)dy$$

and therefore $\sigma(g)f=0$. Hence if we can prove that σ is irreducible (and therefore also quasi-simple (see [10(b)])), our assertion would follow from the Corollary to Lemma 2. However, in view of the above remarks, we can define a linear mapping A of $\mathfrak{H}_{\mathfrak{D}_0}$ onto $\sigma(L(\mathfrak{D}_0))f$ such that

$$A(\pi(g)\psi)=\sigma(g)f \qquad (g\in L(\mathfrak{D}_0)).$$

Then $\dim\sigma(L(\mathfrak{D}_0))f\leqq\dim\mathfrak{H}_{\mathfrak{D}_0}<\infty$ and therefore since $\sigma(L(\mathfrak{D}_0))f$ is dense in $U_{\mathfrak{D}_0}$, $U_{\mathfrak{D}_0}=\sigma(L(\mathfrak{D}_0))f$. But $\mathfrak{H}_{\mathfrak{D}_0}$ is irreducible under $\pi(L(\mathfrak{D}_0))$ and so it follows from the existence of A that the same is true for $U_{\mathfrak{D}_0}$ under $\sigma(L(\mathfrak{D}_0))$. Now suppose V is a closed subspace of U which is invariant under $\sigma(G)$ and let W be the orthogonal complement of V in U. In view of the above irreducibility either V or W must contain $U_{\mathfrak{D}_0}$. Suppose $V\supset U_{\mathfrak{D}_0}$. Then $f\in V$ and therefore $V=U$ from the definition of U. Similarly if $W\supset U_{\mathfrak{D}_0}$, $W=U$. This proves that σ is an irreducible representation and so our lemma follows.

COROLLARY.[1] *Let π be an irreducible unitary representation of G on a Hilbert space $\mathfrak{H}$. Suppose there exist two elements $\phi_0\neq 0$, $\psi_0\neq 0$ in $\mathfrak{H}$ such that*

$$\int_{G^*}|(\phi_0,\pi(x)\psi_0)|^2\,dx^*<\infty.$$

Then there exists a positive real number d_π such that

$$\int_{G^*}|(\phi,\pi(x)\psi)|^2\,dx^*=d_\pi^{-1}\,|\phi|^2\,|\psi|^2$$

for all ϕ, ψ in $\mathfrak{H}$.

Let V and W respectively be the subspaces of $\mathfrak{H}$ consisting of all elements of the form $\pi(f)\phi_0$ and $\pi(f)\psi_0$ $(f\in C_c(G))$. Then V and W are both dense in $\mathfrak{H}$. Since π is irreducible and unitary, it is quasi-simple [10(b)] and therefore $\dim\mathfrak{H}_{\mathfrak{D}}<\infty$ $(\mathfrak{D}\in\Omega)$. But if $\mathfrak{D}\in\Omega_\pi$, $E_{\mathfrak{D}}\pi(f)=\pi({}_{\mathfrak{D}}f)$ $(f\in C_c(G))$ and hence $E_{\mathfrak{D}}V\subset V$. V being dense in $\mathfrak{H}$, it follows that $E_{\mathfrak{D}}V$ is dense in $\mathfrak{H}_{\mathfrak{D}}$ and therefore $E_{\mathfrak{D}}V=\mathfrak{H}_{\mathfrak{D}}$. This shows that $\mathfrak{H}_0=\sum_{\mathfrak{D}}\mathfrak{H}_{\mathfrak{D}}\subset V$. Similarly $\mathfrak{H}_0\subset W$.

If $f,g\in C_c(G)$,

$$(\pi(f)\psi_0,\pi(x)\pi(g)\psi_0)=\int\int f(y)\,(\phi_0,\pi(y^{-1}xz)\psi_0)g(z)\,dydz$$

[1] See Godement [4(b)].

and therefore it is obvious that

$$\int_{G^*} |(\phi, \pi(x)\psi)|^2 \, dx < \infty$$

for $\phi \varepsilon V$ and $\psi \varepsilon W$. Hence without loss of generality we may assume that $\phi_0, \psi_0 \varepsilon \mathfrak{H}_{\mathfrak{D}_0}$ for some $\mathfrak{D}_0 \varepsilon \Omega_\pi$. Let $\phi \neq 0$ be any element in V. Put $f_\phi(x) = (\phi, \pi(x)\psi_0)$ $(x \varepsilon G)$ and define U' and σ' as in the proof of Lemma 3. Let U_ϕ be the smallest closed subspace of U' containing f_ϕ which is invariant under $\sigma'(G)$ and let σ_ϕ denote the corresponding representation of G on U_ϕ. Then as we have seen during the proof of Lemma 3, σ_ϕ is irreducible and quasi-simple. Also it is obvious that $\sigma_\phi(h)f_\phi$ $(h \varepsilon C_c(G))$ is the function $x \to (\phi, \pi(x)\pi(h)\psi_0)$ $(x \varepsilon G)$. Therefore it follows from the Corollary to Lemma 2 that π and σ_ϕ are infinitesimally equivalent. Since they are both unitary they are equivalent [5(b), Theorem 8] and so there exists a unitary mapping B_ϕ of $\mathfrak{H}$ onto U_ϕ such that $B_\phi \pi(x) = \sigma_\phi(x) B_\phi$ $(x \varepsilon G)$. Moreover in view of what we have said above, there exists a linear mapping A_ϕ of W into U_ϕ such that

$$A_\phi \pi(h)\psi_0 = \sigma_\phi(h) f_\phi \qquad (h \varepsilon C_c(G)).$$

Put $C_\phi = B_\phi^{-1} A_\phi$. Then C_ϕ is a linear mapping of W into $\mathfrak{H}$ and $\pi(h)C_\phi = C_\phi \pi(h)$ $(h \varepsilon C_c(G))$. This holds in particular if $h \varepsilon L(\mathfrak{D})$ $(\mathfrak{D} \varepsilon \Omega_\pi)$ and therefore C_ϕ maps $\mathfrak{H}_{\mathfrak{D}}$ into itself. Now if we apply Schur's lemma to the finite-dimensional irreducible representation of $L(\mathfrak{D})$ on $\mathfrak{H}_{\mathfrak{D}}$, we can conclude from Lemma 1 that C_ϕ must be a scalar multiple of the identity on $\mathfrak{H}_0$. So there exists a complex number c_ϕ such that

$$\| \sigma_\phi(h) f_\phi \| = \| A_\phi \pi(h)\psi_0 \| = | c_\phi | \, \| B_\phi \pi(h)\psi_0 \| = | c_\phi | \, | \pi(h)\psi_0 |$$

if $h \varepsilon L_\pi$. (Here $\| \;\; \|$ denotes the norm in U_ϕ.) But since

$$\| \sigma_\phi(h) f_\phi \|^2 = \int_{G^*} |(\phi, \pi(x)\pi(h)\psi_0)|^2 \, dx^*,$$

it follows that

$$\int_{G^*} |(\phi, \pi(x)\psi)|^2 \, dx^* = | c_\phi |^2 \, | \psi |^2$$

for all $\psi \varepsilon \mathfrak{H}_0$. If $\phi = 0$, we put $c_\phi = 0$ so that the above relation continues to hold in that case as well. Now suppose ϕ, ψ lie in $\mathfrak{H}_0 \subset V \cap W$. Then

$$\int |(\phi, \pi(x)\psi)|^2 \, dx^* = \int |(\psi, \pi(x)\phi)|^2 \, dx^*$$

and therefore

$$| c_\phi |^2 \, | \psi |^2 = | c_\psi |^2 \, | \phi |^2.$$

Hence we can find a real number c such that $|c_\phi|^2 = c|\phi|^2$ for $\phi \varepsilon \mathfrak{H}_0$ and therefore

$$\int |(\phi, \pi(x)\psi)|^2 dx^* = c|\phi|^2|\psi|^2 \qquad (\phi, \psi \varepsilon \mathfrak{H}_0).$$

Since $\mathfrak{H}_0$ is dense in $\mathfrak{H}$, it is obvious that c is positive and an elementary argument shows that the above relation continues to hold for all $\phi, \psi \varepsilon \mathfrak{H}$. Now if we put $d_\pi = c^{-1}$ we get the assertion of the Corollary.

3. The Schur orthogonality relations.

DEFINITION. *Let π be an irreducible unitary representation of G on a Hilbert space $\mathfrak{H}$. We say that π is square-integrable, if there exist two elements $\phi_0 \neq 0$, $\psi_0 \neq 0$ in $\mathfrak{H}$ such that*

$$\int_{G^*} |(\phi_0, \pi(x)\psi_0)|^2 dx^* < \infty.$$

Similarly we say that π is integrable if

$$\int_{G^*} |(\phi_0, \pi(x)\psi_0)| dx^* < \infty$$

for some nonzero elements ϕ_0, ψ_0 in $\mathfrak{H}$.

It is obvious that the above definitions do not depend on the choice of the subgroup Z_0 so long as Z/Z_0 is finite. We have seen above (Corollary to Lemma 3) that if π is square-integrable

$$\int_{G^*} |(\phi, \pi(x)\psi)|^2 dx < \infty$$

for all $\phi, \psi \varepsilon \mathfrak{H}$. In analogy with the case of compact groups, we shall call the number d_π (of the Corollary to Lemma 3) the *formal degree* of π. Naturally d_π depends on the choice of Z_0 and the normalization of the Haar measure of G^*. However once these have been fixed, it is obvious that two equivalent square-integrable representations have the same formal degree.

The situation for integrable representations is somewhat similar.

LEMMA 4. *Let π be an integrable representation of G on $\mathfrak{H}$. Then if $\mathfrak{H}_0 = \sum_{\mathfrak{D} \varepsilon \Omega} \mathfrak{H}_{\mathfrak{D}}$,*

$$\int |(\phi, \pi(z)\psi)| dx^* < \infty$$

for all ϕ, ψ in $\mathfrak{H}_0$.

Choose nonzero elements ϕ_0, ψ_0 in $\mathfrak{H}$ such that

$$\int |(\phi_0, \pi(x)\psi_0)| \, dx^* < \infty.$$

Then if $g, h \in C_c(G)$, it is easy to verify that the function $|(\pi(g)\phi_0, \pi(x)\pi(h)\psi_0)|$ is integrable on G^*. Let V and W be the set of all elements of the form $\pi(g)\phi_0$ and $\pi(g)\psi_0$ $(g \in C_c(G))$ respectively. Then, as we have seen during the proof of the Corollary to Lemma 3, $\mathfrak{H}_0 \subset V \cap W$ and therefore our assertion follows.

Let π and π' be two square-integrable representations of G on the Hilbert spaces $\mathfrak{H}$ and $\mathfrak{H}'$ respectively. Then if their central characters coincide on Z_0, it is obvious that[2] $(\phi, \pi(x)\psi) \operatorname{conj}(\phi', \pi'(x)\psi')$ may be regarded as a function of x^* on G^* $(\phi, \psi \in \mathfrak{H}; \phi', \psi' \in \mathfrak{H}'; x \in G)$.

THEOREM 1 (The Schur orthogonality relations[1]). *If π and π' are not equivalent*

$$\int_{G^*} (\phi, \pi(x)\psi) \operatorname{conj}(\phi', \pi'(x)\psi') \, dx^* = 0$$

for all $\phi, \psi \in \mathfrak{H}$ and $\phi', \psi' \in \mathfrak{H}'$. On the other hand if the two representations are equivalent under a unitary mapping U of $\mathfrak{H}$ onto $\mathfrak{H}'$,

$$\int_{G^*} (\phi, \pi(x)\psi) \operatorname{conj}(\phi', \pi'(x)\psi') \, dx^* = d_\pi^{-1} (U\phi, \phi') (\psi', U\psi)$$

$(\phi, \psi \in \mathfrak{H}; \phi' \psi' \in \mathfrak{H}')$ *where d_π is the formal degree of π.*

Let $d_{\pi'}$ denote the formal degree of π'. Then it is obvious from the Corollary to Lemma 3 that

$$\int |(\phi, \pi(x)\psi) \operatorname{conj}(\phi', \pi'(x)\psi')| \, dx^* \leqq (d_\pi d_{\pi'})^{-1} |\phi|^2 |\psi|^2 |\phi'|^2 |\psi'|^2.$$

Therefore for any given $\phi \in \mathfrak{H}$ and $\phi' \in \mathfrak{H}'$, there exists a bounded linear operator A from $\mathfrak{H}$ to $\mathfrak{H}'$ such that

$$(\psi', A\psi) = \int (\phi, \pi(x)\psi) \operatorname{conj}(\phi', \pi'(x)\psi') \, dx^*$$

for all $\psi' \in \mathfrak{H}'$ and $\psi \in \mathfrak{H}$. It follows immediately from this relation that

$$(\psi', A\pi(x)\psi) = (\pi'(x^{-1})\psi', A\psi) \qquad (x \in G)$$

and therefore $A\pi(x) = \pi'(x)A$. In order to prove the first statement of the

---

[2] conj c denotes the conjugate of a complex number c.

theorem it would be enough to show that if $A \neq 0$, π and π' are equivalent. So let us suppose $A \neq 0$. Choose ψ in $\mathfrak{H}$ such that $A\psi \neq 0$. Then if $h \varepsilon C_c(G)$ it is clear that

$$A\pi(h)\psi = \pi'(h)A\psi$$

and therefore from the Corollary to Lemma 2, π and π' are infinitesimally equivalent. But since they are both unitary this implies that they are equivalent [5(b), Theorem 8].

In order to prove the second statement we may assume that $\pi' = \pi$ since

$$(\phi', \pi'(x)\psi') = (U^{-1}\phi', \pi(x)U^{-1}\psi')$$

in the general case. Hence if we keep to the above notation, A is now a bounded linear operator on $\mathfrak{H}$, which commutes with $\pi(x)$ $(x \varepsilon G)$. Since π is irreducible, A must be a scalar multiple of the identity. Hence

$$\int (\phi, \pi(x)\psi)\operatorname{conj}(\phi', \pi(x)\psi')\,dx^* = c_{\phi,\phi'}(\psi', \psi)$$

where $c_{\phi,\phi'}$ is a complex number depending only on ϕ and ϕ'. But obviously

$$\int (\phi, \pi(x)\psi)\operatorname{conj}(\phi', \pi(x)\psi')\,dx^* = \int (\psi', \pi(x)\phi')\operatorname{conj}(\psi, \pi(x)\phi)\,dx^*$$

and therefore

$$c_{\phi,\phi'}(\psi', \psi) = c_{\psi,\psi'}(\phi, \phi').$$

Choose $\psi' = \psi = \psi_0 \neq 0$ and put $c = (\psi_0, \psi_0)/|\psi_0|^2$. Then $c_{\phi,\phi'} = c(\phi', \phi)$ for all $\phi', \phi \varepsilon \mathfrak{H}$. In particular if we put $\phi = \phi' = \psi = \psi' = \psi_0$ we find that $c = d_\pi^{-1}$. Thus the theorem is proved.

4. The character of a square-integrable representation. Let $C_c^\infty(G)$ denote, as before, the subspace of $C_c(G)$ consisting of those functions which are indefinitely differentiable everywhere. For $x^* \varepsilon G^*$, we put $y^{x^*} = xyx^{-1}$ $(y \varepsilon G)$ where x is any element in G whose image in G^* is x^*.

THEOREM 2. *Let π be a square-integrable representation of G on a Hilbert space and let T_π denote the character* [5(c)] *of π. Then if $f \varepsilon C_c^\infty(G)$,*

$$T_\pi(f) = d_\pi \int_{G^*} dx^* \{ \int_G f(y^{x^*})(\phi, \pi(y)\phi)\,dy \}$$

where ϕ is any unit vector in $\mathfrak{H}$ and d_π is the formal degree of π.

Let Q denote the operator $\int_G f(y)\pi(y)\,dy$. We know [5(c), p. 243] that

there exists a complete orthonormal set $\{\psi_j\}_{j\in J}$ in $\mathfrak{H}$ such that $\sum_{i,j\in J}|Q_{ij}|<\infty$ where $Q_{ij}=(\psi_i,Q\psi_j)$. Moreover $\pi(x^{-1})Q\pi(x)$ $(x\in G)$ depends only on x^* and so we may denote it by Q^{x^*}. Then

$$\begin{aligned}(\phi,Q^{x^*}\phi)=(\pi(x)\phi,Q\pi(x)\phi)&=\sum_i(\pi(x)\phi,\psi_i)(\psi_i,Q\pi(x)\phi)\\&=\sum_i\sum_j(\pi(x)\phi,\psi_i)Q_{ij}(\psi_j,\pi(x)\phi).\end{aligned}$$

But if we make use of the Schwartz inequality and the Schur orthogonality relations, we get

$$\begin{aligned}&\sum_{i,j}\int_{G^*}|(\pi(x)\phi,\psi_i)Q_{ij}(\psi_j,\pi(x)\phi)|\,dx^*\\&\quad\leqq\sum_{i,j}|Q_{ij}|\{\int_{G^*}|(\pi(x)\phi,\psi_i)|^2dx^*\int_{G^*}|(\psi_j,\ (x)\phi)|^2dx^*\}^{\frac12}\\&\quad=d_\pi^{-1}\sum_{i,j}|Q_{ij}|<\infty\end{aligned}$$

and therefore by Lebegue's Theorem the above series for $(\phi,Q^{x^*}\phi)$ may be integrated over G^* term by term. Hence

$$\begin{aligned}\int_{G^*}(\phi,Q^{x^*}\phi)\,dx^*&=\sum_i\sum_jQ_{ij}\int_{G^*}(\pi(x)\phi,\psi_i)(\psi_j,\pi(x)\phi)\,dx^*\\&=\sum_i\sum_j(Q_{ij}/d_\pi)(\psi_j,\psi_i)=d_\pi^{-1}\sum Q_{ii}\\&=d_\pi^{-1}\operatorname{Sp}Q=d_\pi^{-1}T_\pi(f).\end{aligned}$$

But

$$(\phi,Q^{x^*}\phi)=\int_G f(y^{x^*})(\phi,\pi(y)\phi)\,dy$$

and so the theorem is proved.

It should be noticed that, in general, the double integral in the above theorem is not absolutely convergent and therefore the order of the two integrations cannot be interchanged.

5. The discrete part of the Plancherel measure. We shall assume in this section that Z is finite and $Z_0=\{1\}$. Let $\mathcal{E}$ denote the set of all equivalence classes of irreducible unitary representations of G. We consider the Hilbert space $L_2(G)$ consisting of all complex-valued functions on G which are square-integrable with respect to the Haar measure. Let λ denote the left regular representation of G on $L_2(G)$ defined by

$$(\lambda(x)f)(y)=f(x^{-1}y)\qquad\qquad(x,y\in G;f\in L_2(G)).$$

Then λ is unitary. We say that a class $\omega \varepsilon \mathcal{E}$ is discrete if there exists a closed subspace $\mathfrak{H} \neq 0$ of $L_2(G)$ which is invariant and irreducible under $\lambda(G)$ and such that the corresponding representation of G on $\mathfrak{H}$ lies in ω. It is known (see Godemont [4(a), Theorem 1]) that ω is discrete if and only if every representation in ω is square-integrable. Let $\mathcal{E}_0$ denote the set of all discrete classes in $\mathcal{E}$. If $\omega \varepsilon \mathcal{E}_0$, we denote by d_ω the formal degree of any representation in ω.

For any $\omega \varepsilon \mathcal{E}$, let T_ω denote the character [5(c)] of any representation in ω. Then it is known (see Segal [10(b)], Mautner [8] and [5(b), Theorem 7]) that there exists a unique positive measure μ on $\mathcal{E}$ such that

$$\int_G |f(x)|^2 \, dx = \int_{\mathcal{E}} T_\omega(\tilde{f} * f)\, d\mu \qquad (f \varepsilon C_c(G))$$

where $\tilde{f}(x) = \operatorname{conj} f(x^{-1})$ $(x \varepsilon G)$. We shall call μ the Plancherel measure on $\mathcal{E}$. First we prove the following simple lemma.

LEMMA 5. *Every single point ω_0 in $\mathcal{E}$ is μ-measurable.*

For any $f \varepsilon C_c(G)$ put $\| f \|_1 = \int_G | f(x) | \, dx$. Then under this norm $C_c(G)$ becomes a separable metric space. Let π_0 be a representation in ω_0 and let $\mathfrak{M}_{\omega_0}$ denotes the set of all $f \varepsilon C_c(G)$ such that $\pi_0(f) = 0$. Then $\mathfrak{M}_{\omega_0}$ is also separable under the above metric and so we can select a sequence $\{\alpha_1, \alpha_2, \cdots\}$ in $\mathfrak{M}_{\omega_0}$ which is dense in $\mathfrak{M}_{\omega_0}$. Put $F_n(\omega) = T_\omega(\tilde{\alpha}_n * \alpha_n)$ $(\omega \varepsilon \mathcal{E})$. Then $F_1, F_2, \cdots$ are all measurable functions on $\mathcal{E}$ and it follows from the Corollary to Lemma 2 that ω_0 is the only point in $\mathcal{E}$ where they vanish simultaneously. From this the lemma follows immediately.

Our main object in this section is to prove the following theorem.

THEOREM 3. *If ω_0 is a discrete class, $\mu(\omega_0) = d_{\omega_0}$.*

Define π_0, $\mathfrak{M}_{\omega_0}$ and $\{\alpha_n\}_{n \geqq 1}$ as above and let ψ_0 be a nonzero vector in the representation space $\mathfrak{H}$ of π_0. Put $g(x) = (\pi_0(x)\psi_0, \psi_0)$. Then since ω_0 is discrete, $g \varepsilon L_2(G)$. We first need the following lemma.

LEMMA 6. *There exists a sequence $g_1, g_2, \cdots$ in $C_c(G)$ and a subset $\mathcal{E}'$ of $\mathcal{E}$ satisfying the following conditions:*

(1) *The complement of $\mathcal{E}'$ in $\mathcal{E}$ is of μ-measure zero.*

(2) $\lim_{n \to \infty} g_n = g$ *in* $L_2(G)$.

(3) *For every* $\omega \in \mathcal{E}'$, $\operatorname*{Lim}_{n\to\infty} T_\omega\ (\tilde{g}_n * g_n)$ *exists and is finite and*[3] $\operatorname*{Lim}_{n\to\infty} T_\omega((\alpha_m * g_n)^{\sim} * (\alpha * g_n)) = 0$ *for every* $m \geqq 1$.

Since $C_c(G)$ is dense in $L_2(G)$ we can choose a sequence $\{h_1, h_2, \cdots\}$ in $C_c(G)$ such that $h_n \to g$ in $L_2(G)$. Then

$$\int_{\mathcal{E}} T_\omega(\bar{h}_n * h_n)\, d\mu = \| h_n \|^2 \to \| g \|^2$$

where $\|\ \|$ denotes the norm in $L_2(G)$. Therefore by the Riesz-Fischer Theorem, there exists a subset $\mathcal{E}_0' \subset \mathcal{E}$ and a subsequence $\{h_n^{(0)}\}$ of $\{h_n\}$ such that[4] (1) $\mathcal{E} - \mathcal{E}_0'$ is of μ-measure zero and (2) $\operatorname*{Lim}_{n\to\infty} T\ (h_n^{(0)\sim} * h_n^{(0)})$ exists and is finite for every $\omega \in \mathcal{E}_0'$. Now for each integer $r \geqq 0$ we shall define a subsequence $\{h_n^{(r)}\}$ of $\{h_n^{(0)}\}$ and a subset $\mathcal{E}_r' \subset \mathcal{E}_0'$ such that (1) $\mathcal{E} - \mathcal{E}_r'$ is of μ-measure zero and (2) $\operatorname*{Lim}_{n\to\infty} T_\omega((\alpha_m * h_n^{(r)})^{\sim} * (\alpha_m * h_n^{(r)})) = 0$ for $\omega \in \mathcal{E}_r'$ and $1 \leqq m \leqq r$. This has already been done for $r = 0$. So assuming that $\{h_n^{(r)}\}$ and $\mathcal{E}_r'$ have been defined, we proceed to define $\{h_n^{(r+1)}\}$ and $\mathcal{E}_{r+1}'$ by induction. Suppose $\omega \in \mathcal{E}_r'$. Then $T_\omega(h_n^{(0)\sim} * h_n^{(0)})$ converges to a finite limit and therefore if $\pi \in \omega$, $\pi(h_n^{(r)})$ converges with respect to the Hilbert-Schmidt (H. S.) norm to a bounded operator A_π. Now put $h_n' = \alpha_{r+1} * h_n^{(r)}$. Then $\pi(h_n') = \pi(\alpha_{r+1})\pi(h_n^{(r)})$ and since $\pi(\alpha_{r+1})$ is a bounded operator, it is obvious that $\pi(h_n')$ also converges to $\pi(\alpha_{r+1})A_\pi$ in the H. S. norm. On the other hand since $\| h_n^{(r)} - g \| \to 0$, $\alpha_{r+1} * h_n^{(r)} \to \alpha_{r+1} * g$ in $L_2(G)$. But

$$(\alpha_{r+1} * g)(x) = (\pi_0(x)\psi_0, \pi_0(\alpha_{r+1})\psi_0) = 0 \qquad (x \in G)$$

because $\alpha_{r+1} \in \mathfrak{M}_{\omega_0}$. Therefore $\alpha_{r+1} * h_n^{(r)} \to 0$ in $L_2(G)$ and so

$$\int_{\mathcal{E}} T_\omega((\alpha_{r+1} * h_n^{(r)})^{\sim} * (\alpha_{r+1} * h_n^{(r)}))\, d\mu \to 0$$

as $n \to \infty$. Hence by the Riesz-Fischer Theorem we can select a subsequence $\{h_n^{(r+1)}\}$ of $\{h_n^{(r)}\}$ and a subset $\mathcal{E}_{r+1} \subset \mathcal{E}$ such that (1) $\mathcal{E} - \mathcal{E}_{r+1}$ is of μ-measure zero and (2) if $\omega \in \mathcal{E}_{r+1}$,

$$\operatorname*{Lim}_{n\to\infty} T_\omega((\alpha_{r+1} * h_n^{(r+1)})^{\sim} * (\alpha_{r+1} * h_n^{(r+1)})) = 0.$$

Now put $\mathcal{E}_{r+1}' = \mathcal{E}_r' \cap \mathcal{E}_{r+1}$. Then $\mathcal{E}_{r+1}$ and $\{h_n^{(r+1)}\}$ satisfy all the required conditions. Hence if we take $\mathcal{E}' = \bigcap_{r=1}^{\infty} \mathcal{E}_r'$ and $g_n = h_n^{(n)}$ we get the assertion of the lemma.

---

[3] We write $f^{\sim}$ instead of $\bar{f}$ $(f \in C_c(G))$ whenever it is convenient to do so.

[4] $\mathcal{E} - \mathcal{E}_0'$ is the complement of $\mathcal{E}_0'$ in $\mathcal{E}$.

Let us now come to the proof of the theorem. Let ω be a class in $\mathcal{E}'$ and π a representation in ω. Since $T_\omega(\tilde{g}_n * g_n)$ is convergent, the operators $\pi(g_n)$ converge to a limit in the H. S. norm. Let A_π denote this limit. Since $\pi(\alpha_r)$ is a bounded operator $\pi(\alpha_r * g_n) = \pi(\alpha_r)\pi(g_n)$ also converges to $\pi(\alpha_r)A_\pi$ in the H. S. norm. However

$$\lim_{n\to\infty} T_\omega((\alpha_r * g_n)^{\sim} * (\alpha_r * g_n)) = 0$$

and therefore $\pi(\alpha_r)A_\pi = 0$ $(r \geqq 1)$. Now suppose $\omega \neq \omega_0$. Then if ψ lies in the representation space of π, $\pi(\alpha_r)A_\pi\psi = 0$. Since π and π_0 are not equivalent (and therefore also not infinitesimally equivalent [5(b), Theorem 8]), it follows from the Corollary to Lemma 2 that $A_\pi\psi = 0$. This being true for every ψ, $A_\pi = 0$. But $\pi(g_n)$ tends to A_π in the H. S. norm and so this shows that

$$\lim_{n\to\infty} T_\omega(\tilde{g}_n * g_n) = 0.$$

On the other hand $g_n \to g$ in $L_2(G)$ and therefore

$$\| g \|^2 = \lim_{n\to\infty} \int_{\mathcal{E}} T_\omega(\tilde{g}_n * g_n)\, d\mu.$$

From this it follows that $\omega_0 \in \mathcal{E}'$. For otherwise, in view of what we have just said, $\lim_{n\to\infty} T_\omega(\tilde{g}_n * g_n) = 0$ for all $\omega \in \mathcal{E}'$ and since $\mathcal{E} - \mathcal{E}'$ is of μ-measure zero we would have

$$\| g \|^2 = \lim_{n\to\infty} \int_{\mathcal{E}} T_\omega(\tilde{g}_n * g_n)\, d\mu = 0.$$

But this is false since g is continuous and $g(1) = |\psi_0|^2 \neq 0$. Therefore $\omega_0 \in \mathcal{E}'$ and so $T_{\omega_0}(\tilde{g}_n * g_n)$ tends to a finite limit. This shows that

$$\begin{aligned} \| g \|^2 &= \lim_{n\to\infty} \int_{\mathcal{E}'} T_\omega(\tilde{g}_n * g_n)\, d\mu = \int_{\mathcal{E}'} \lim_{n\to\infty} T_\omega(\tilde{g}_n * g_n)\, d\mu \\ &= \mu(\omega_0) \lim_{n\to\infty} T_{\omega_0}(\tilde{g}_n * g_n). \end{aligned}$$

Now let[5] $\{\psi_1, \psi_2, \cdots\}$ be a complete orthonormal set in the representation space $\mathfrak{H}$ of π_0. Then if $A_n = \pi_0(g_n)$,

$$(\psi_i, A_n\psi_j) = \int_G g_n(x)\,(\psi_i, \pi_0(x)\psi_j)\, dx.$$

Since π_0 is square-integrable it follows that

$$\lim_{n\to\infty} (\psi_i, A_n\psi_j) = \int g(x)(\psi_i, \pi_0(x)\psi_j) dx = (\psi_i, \pi_0(g)\psi_j).$$

---

[5] It is not difficult to see that $\mathfrak{H}$ is separable.

But we have seen above that A_n converges in the H. S. norm and so its limit must be $\pi_0(g)$. This proves that

$$\operatorname*{Lim}_{n\to\infty} T_{\omega_0}(\tilde{g}_n * g_n) = T_{\omega_0}(\tilde{g} * g)$$

and therefore

$$\| g \|^2 = \mu(\omega_0) T_{\omega_0}(\tilde{g} * g).$$

But

$$(\psi_i, \pi_0(g)\psi_j) = \int (\pi_0(x)\psi_0, \psi_0)(\psi_i, \pi_0(x)\psi_j)\,dx$$
$$= d_{\omega_0}^{-1}(\psi_0, \psi_j)(\psi_i, \psi_0).$$

Hence

$$T_{\omega_0}(\tilde{g} * g) = \sum_{i,j} |(\psi_i, \pi_0(g)\psi_j)|^2 = d_{\omega_0}^{-2} \sum_{i,j} |(\psi_0, \psi_j)(\psi_i, \psi_0)|^2 = d_{\omega_0}^{-2} |\psi_0|^4.$$

On the other hand

$$\| g \|^2 = \int |(\pi(x)\psi_0, \psi_0)|^2\,dx = d_{\omega_0}^{-1} |\psi_0|^4$$

and therefore $\mu(\omega_0) = d_{\omega_0}$.

COROLLARY 1. *Let π be a square-integrable representation of G. Then if $f \varepsilon C_c(G)$,*

$$\left\| \int f(x)\pi(x)\,dx \right\|^2 \leqq d_\pi^{-1} \int |f(x)|^2\,dx$$

where $\| A \|$ denotes the H. S. *norm of an operator A.*

For $\int |f(x)|^2\,dx = \int_{\mathcal{E}} T_\omega(\tilde{f} * f)\,d\mu \geqq \mu(\omega_0) T_{\omega_0}(\tilde{f} * f)$ if ω_0 is the class of π. But $\mu(\omega_0) = d_\pi$ and

$$T_{\omega_0}(\tilde{f} * f) = \left\| \int f(x)\pi(x)\,dx \right\|^2.$$

Hence the result

COROLLARY 2. *A class $\omega_0 \varepsilon \mathcal{E}$ is discrete if and only if $\mu(\omega_0) > 0$.*

We have seen that if ω_0 is discrete $\mu(\omega_0) = d_{\omega_0}$ is positive. Conversely suppose $\mu(\omega_0)$ is positive. Let π be a representation in ω_0 and ϕ a unit vector in the representation space of π. Then if $f \varepsilon C_c(G)$,

$$|(\phi, \pi(f)\phi)|^2 \leqq \left\| \int f(x)\pi(x)\,dx \right\|^2 = T_{\omega_0}(f * f).$$

On the other hand

$$\| f \|^2 = \int_{\mathcal{E}} T_\omega(\tilde{f} * f)\,d\mu \geqq \mu(\omega_0) T_{\omega_0}(\tilde{f} * f).$$

Therefore

$$|(\phi, \pi(f)\phi)|^2 \leqq \mu(\omega_0)^{-1} \| f \|^2.$$

Let $L_1(G)$ denote, as usual, the space of all functions which are integrable on G. Then if $L = L_1(G) \cap L_2(G)$, it follows from the above inequality that

$$|(\phi, \pi(g)\phi)|^2 \leqq \mu(\omega_0)^{-1} \| g \|^2 \qquad (g \varepsilon L)$$

where $\pi(g) = \int g(x)\pi(x)\,dx$. U being any compact neighborhood of 1 in G, we now define a function $g_U \varepsilon L$ as follows. $g_U(x) = (\pi(x)\phi, \phi)$ if $x \varepsilon U$ and $g_U(x) = 0$ otherwise. Then

$$(\phi, \pi(g_U)\phi) = \int g_U(x)\,(\phi, \pi(x)\phi)\,dx = \int_U |(\phi, \pi(x)\phi)|^2\,dx = \| g_U \|^2$$

and therefore

$$\| g_U \|^2 = \int_U |(\phi, \pi(x)\phi)|^2\,dx \leqq \mu(\omega_0)^{-1}.$$

Hence

$$\int |(\phi, \pi(x)\phi)|^2\,dx = \sup_U \int_U |(\phi, \pi(x)\phi)|^2\,dx \leqq \mu(\omega_0)^{-1}.$$

This proves that π is square-integrable and therefore ω_0 is discrete.

Part II.

6. Some algebraic results. We shall now study in detail certain special representations which have been constructed in another paper [5(f)] and prove that, under suitable conditions, they are square-integrable or even integrable. Later (in Sections 9 and 10) we shall also obtain a formula for the formal degree of these representations.

Let $\mathfrak{h}_0$ be a maximal abelian subalgebra of $\mathfrak{k}_0$. In accordance with the assumption of [5(e), (f)] *we shall suppose that* $\mathfrak{h}_0$ *is also maximal abelian in* $\mathfrak{g}_0$. From now on we use the notation and the terminology of [5(e), § 3] without further comment. Then $\mathfrak{h}$ is a Cartan subalgebra of $\mathfrak{g}$. Suppose an order has been introduced once for all in the space $\mathfrak{F}_R$ of real linear functions on $\mathfrak{h}$ (see [5(e), § 2]) and P is the set of all positive roots of $\mathfrak{g}$ (with respect to $\mathfrak{h}$) in this order. *We shall further assume that every noncompact root is totally positive.* Since we are now interested primarily in unitary representations, this assumption is justified in view of Corollary 1 to Lemma 19 of [5(e)].

Let $\mathfrak{k}$ be a subalgebra of $\mathfrak{g}$. Suppose there exists a set Q of totally positive roots such that

$$\mathfrak{l} = \mathfrak{l} \cap \mathfrak{k} + \sum_{\gamma \varepsilon Q} (CX_\gamma + CX_{-\gamma})$$

and let β be the lowest root in Q. Then X_β, $X_{-\beta}$ and therefore also $H_\beta = [X_\beta, X_{-\beta}]$ are in $\mathfrak{l}$. Let $\mathfrak{l}_\beta$ denote the centralizer of $CH_\beta + CX_\beta + CX_{-\beta}$ in $\mathfrak{l}$. It is obvious that $\mathfrak{l}_\beta$ is invariant under θ and therefore

$$\mathfrak{l}_\beta = \mathfrak{l}_\beta \cap \mathfrak{k} + \mathfrak{l}_\beta \cap \mathfrak{p}.$$

LEMMA 7. *$C(X_\beta + X_{-\beta}) + \mathfrak{l}_\beta \cap \mathfrak{p}$ is exactly the set of all elements in $\mathfrak{l} \cap \mathfrak{p}$ which commute with $X_\beta + X_{-\beta}$.*

Let Q' be the set of all roots in Q other than β. Then if $X \varepsilon \mathfrak{l} \cap \mathfrak{p}$,

$$X = c_\beta' X_\beta + c_{-\beta}' X_{-\beta} + \sum_{\gamma \varepsilon Q'} (c_\gamma X + c_{-\gamma} X_{-\gamma})$$

where c_β', $c_{-\beta}'$, c_γ, $c_{-\gamma}$ are complex numbers. Now

$$\mathfrak{g} = \mathfrak{h} + \sum_{\delta \varepsilon P} (CX_\delta + CX_{-\delta})$$

where the sum is direct. Hence it is clear that the component of $[X, X_\beta + X_{-\beta}]$ in $\mathfrak{h}$ is $(c_\beta' - c_{-\beta}')H_\beta$. So if X commutes with $X_\beta + X_{-\beta}$, $c_\beta' = c_{-\beta}'$ and therefore

$$Y = \sum_{\gamma \varepsilon Q} (c_\gamma X_\gamma + c_{-\gamma} X_{-\gamma})$$

also commutes with $(X_\beta + X_{-\beta})$. In order to prove the lemma it is enough to show that $Y \varepsilon \mathfrak{l}_\beta$. Let us suppose then that this is false. Define $c_\delta = 0$ for any root δ for which neither δ nor $-\delta$ is in Q'. Then it is obvious that there exists roots δ such that (1) $c_\delta \neq 0$ and (2) $X_\delta \notin \mathfrak{l}_\beta$, for otherwise Y would lie in $\mathfrak{l}_\beta$. Let δ_0 be the highest such root. Since $[Y, X_\beta + X_{-\beta}] = 0$, it follows that $\delta_0 + \beta$ is not a root. However $X_{\delta_0} \notin \mathfrak{l}_\beta$ and so $\delta_0 - \beta$ must be a root. The coefficient of $X_{\delta_0 - \beta}$ in $[Y, X_\beta]$ (with respect to the above decomposition of $\mathfrak{g}$ as a direct sum) is then clearly different from zero. This means that $\delta_0 - \beta = \gamma + \beta$ where γ is some root with $c_\gamma \neq 0$. Hence $\gamma = \delta_0 - 2\beta$ is a root and $X_{\delta_0 - 2\beta} \varepsilon \mathfrak{l}$. Since δ_0 and β are both noncompact, $\alpha = \delta_0 - \beta$ is compact. Moreover β being totally positive, $\delta_0 = \beta + \alpha$ and $2\beta - \delta_0 = \beta - \alpha$ are also totally positive (Lemma 12 of [5(e)]) and $X_{\beta+\alpha} = X_{\delta_0}$, $X_{\beta-\alpha} = X_{2\beta - \delta_0}$ are both in $\mathfrak{l}$. Therefore $\beta + \alpha$ and $\beta - \alpha$ are in Q. This however is impossible since β is the lowest root in Q and so the lemma is proved.

Let Q_β be the set of all $\gamma \varepsilon Q$ such that $\gamma \neq \beta$ and neither $\gamma + \beta$ nor $\gamma - \beta$ is a root. Then it is obvious that

$$\mathfrak{l}_\beta = \mathfrak{l}_\beta \cap \mathfrak{k} + \sum_{\gamma \varepsilon Q_\beta} (CX_\gamma + CX_{-\gamma}).$$

Therefore $\mathfrak{l}_\beta$ satisfies the same condition as the one imposed above on $\mathfrak{l}$. Now we shall define a sequence $\mathfrak{g} = \mathfrak{g}_1 \supset \mathfrak{g}_2 \supset \mathfrak{g}_3 \supset \cdots$ of subalgebras of $\mathfrak{g}$ such that each $\mathfrak{g}_r$ satisfies this condition. The inductive definition is as follows. If $\mathfrak{g}_r \subset \mathfrak{k}$, $\mathfrak{g}_{r+1} = \mathfrak{g}_r$. Otherwise let β be the lowest totally positive root such that $X_\beta \varepsilon \mathfrak{g}_r$. Then $\mathfrak{g}_{r+1}$ is the centralizer of $CH_\beta + CX_\beta + CX_{-\beta}$ in $\mathfrak{g}_r$. It is obvious that $\dim \mathfrak{g}_{r+1} < \dim \mathfrak{g}_r$ unless $\mathfrak{g}_r \subset \mathfrak{k}$ and therefore $\mathfrak{g}_r \subset \mathfrak{k}$ if r is sufficiently large. Let $s \geqq 0$ be the least integer such that $\mathfrak{g}_{s+1} \subset \mathfrak{k}$ and let γ_r be the lowest totally positive root such that $X_{\gamma_r} \varepsilon \mathfrak{g}_r$ $(1 \leqq r \leqq s)$. One proves easily by induction on r that if $X_\alpha \varepsilon \mathfrak{g}_r$ for some root α then $X_{-\alpha}$ is also in $\mathfrak{g}_r$.

LEMMA 8. *$\gamma_i \pm \gamma_j$ $(1 \leqq i < j \leqq s)$ is never a root or zero and the elements $(X_{\gamma_i} + X_{-\gamma_i})$ $i = 1, 2, \cdots, s$ span a maximal abelian subspace of $\mathfrak{p}$ over C.*

If $i < j$, $\mathfrak{g}_{i+1} \supset \mathfrak{g}_j$ and therefore $\mathfrak{g}_j$ commutes with X_{γ_i} and $X_{-\gamma_i}$. Hence $\gamma_i \pm \gamma_j$ is not a root or zero. Let $\mathfrak{a}_\mathfrak{p}$ be the subspace of $\mathfrak{p}$ spanned by $(X_{\gamma_i} + X_{-\gamma_i})$ $i = 1, 2, \cdots, s$. Then $\mathfrak{a}_\mathfrak{p}$ is obviously abelian. Let X be an element in $\mathfrak{p}$ which commutes with $\mathfrak{a}_\mathfrak{p}$. We have to show that $X \varepsilon \mathfrak{a}_\mathfrak{p}$. Suppose this is false. Then it is obvious that $X \not\varepsilon \mathfrak{k} + \mathfrak{a}_\mathfrak{p}$. Since $\mathfrak{g}_{s+1} \subset \mathfrak{k}$, we can choose r $(1 \leqq r \leqq s)$ such that $X \varepsilon \mathfrak{g}_r + \mathfrak{a}_\mathfrak{p}$ but $X \not\varepsilon \mathfrak{g}_{r+1} + \mathfrak{a}_\mathfrak{p}$. Let $X = Y + Z$ $(Y \varepsilon \mathfrak{g}_r, Z \varepsilon \mathfrak{a}_\mathfrak{p})$. Since X commutes with $X_{\gamma_r} + X_{-\gamma_r}$, the same holds for Y. Also $Y = X - Z \varepsilon \mathfrak{g}_r \cap \mathfrak{p}$. Therefore we conclude from Lemma 7 that

$$Y = c(X_{\gamma_r} + X_{-\gamma_r}) + Y_1$$

where $Y_1 \varepsilon \mathfrak{g}_{r+1} \cap \mathfrak{p}$ and $c \varepsilon C$. Then $Z_1 = Z + c(X_{\gamma_r} + X_{-\gamma_r})$ lies in $\mathfrak{a}_\mathfrak{p}$ and so

$$X = Y_1 + Z_1 \varepsilon \mathfrak{g}_{r+1} + \mathfrak{a}_\mathfrak{p}.$$

Since this contradicts the definition of r, the lemma follows.

COROLLARY. *Let $\mathfrak{a}_{\mathfrak{p}_0} = \sum_{i=1}^{s} R(X_{\gamma_i} + X_{-\gamma_i})$. Then $\mathfrak{a}_{\mathfrak{p}_0} = \mathfrak{p}_0 \cap \mathfrak{a}_\mathfrak{p}$ and therefore it is a maximal abelian subspace of $\mathfrak{p}_0$.*

We know that $\tilde{\theta}(X_\gamma) = -X_{-\gamma}$ for any root γ (see [5(e, § 4]. There is a small mistake on p. 757 of [5(e)]. In line 22 $\tilde{\theta}$ should be replaced by η

which is the conjugation of $\mathfrak{g}$ with respect to $\mathfrak{g}_0$.). Hence $X_{\gamma_i}+X_{-\gamma_i}\,\varepsilon\,\mathfrak{p}_0$. Moreover if

$$X=\sum_{i=1}^{s} c(X_{\gamma_i}+X_{-\gamma_i})\,\varepsilon\,\mathfrak{p}_0 \qquad (c_i\,\varepsilon\, C),$$

$X=-\bar{\theta}(X)$ and therefore $c_i\,\varepsilon\, R$.

We now need some simple facts about a three dimensional Lie algebra.

LEMMA 9. *Let $\mathfrak{l}$ be the Lie algebra of dimension 3 spanned over C by the elements H, X, Y satisfying the following relations:*

$$[X,Y]=H, \qquad [H,X]=2X, \qquad [H,Y]=-2Y.$$

Let ν denote the automorphism of $\mathfrak{l}$ given by

$$\nu(Z)=\exp\tfrac{\pi}{4}\operatorname{ad}(X-Y))Z \qquad (Z\,\varepsilon\,\mathfrak{l}).$$

Then $\nu(H)=-(X+Y)$, $\nu(X+Y)=H$, $\nu(X-Y)=X-Y$. Moreover if L is any complex analytic group with the Lie algebra $\mathfrak{l}$,

$$\exp t(X+Y)=\exp(zY)\exp(\log(\cosh t)H)\exp zX \qquad (t\,\varepsilon\, C, \cosh t\neq 0)$$

where[6] $z=\tanh t$.

It is well known that $\mathfrak{l}$ is isomorphic to the Lie algebra of the group of all 2×2 complex matrices with determinant 1. Since this group is simply connected, it is enough to prove the above relations in it. Therefore we may identify X, Y, H with matrices as follows:

$$X=\begin{pmatrix}0&1\\0&0\end{pmatrix},\quad Y=\begin{pmatrix}0&0\\1&0\end{pmatrix},\quad H=\begin{pmatrix}1&0\\0&-1\end{pmatrix}.$$

The required relations are now verified by a simple calculation.

Let P_+ be all the totally positive roots of $\mathfrak{g}$. Then $\gamma_i\,\varepsilon\, P_+$ $1\leqq i\leqq s$. Consider the automorphism ν of $\mathfrak{g}$ given by $\nu=\exp\frac{\pi}{4}\operatorname{ad}(\sum_{j=1}^{s}(X_{\gamma_i}-X_{-\gamma_i}))$. It follows from Lemmas 8 and 9 that $\nu(X_{\gamma_j}+X_{-\gamma_j})=H_{\gamma_j}$ and therefore $\nu(\mathfrak{a}_{\mathfrak{p}})=\sum_{i=1}^{s} CH_{\gamma_i}$. This shows that H_{γ_i} and therefore also γ_i $(1\leqq i\leqq s)$ are

[6] Our result is valid with any determination of the logarithm. But for the sake of definiteness let us make the following convention. Choose a fixed square root of -1 in C and denote it by $(-1)^{\frac{1}{2}}$. Then if z is a non-zero complex number

$$\log z=\log|z|+(-1)^{\frac{1}{2}}\phi$$

where $\log|z|$ and ϕ are real and $0\leqq\phi<2\pi$. Hence in particular if z is real and positive $\log z$ is real.

linearly independent. Let $\mathfrak{a}_{\mathfrak{l}}$ be the orthogonal complement of ν $(\mathfrak{a}_{\mathfrak{p}})$ in $\mathfrak{h}$ with respect to the positive definition Hermitian form—$B(\bar{\theta}(X), X)$ $(X \varepsilon \mathfrak{g}$, see [5(e), § 4]). Since $\bar{\theta}(H_\gamma) = -H_\gamma$ for every root γ, it is obvious that $B(H_{\gamma_i}, H) = 0$ and therefore $\gamma_i(H) = 0$ $1 \leqq i \leqq s$ if $H \varepsilon \mathfrak{a}_{\mathfrak{l}}$. This means that $X_{\gamma_i} - X_{-\gamma_i}$ $1 \leqq i \leqq s$ commute with H and therefore $\nu(H) = H$ $(H \varepsilon \mathfrak{a}_{\mathfrak{l}})$. Hence if $\mathfrak{a} = \mathfrak{a}_{\mathfrak{p}} + \mathfrak{a}_{\mathfrak{l}}$, $\nu(\mathfrak{a}) = \nu(\mathfrak{a}_{\mathfrak{p}}) + \mathfrak{a}_{\mathfrak{l}} = \mathfrak{h}$. As ν is an automorphism, it follows that $\mathfrak{a}$ is a Cartan subalgebra of $\mathfrak{g}$.

Let α, β be two roots of $\mathfrak{g}$ and let k, k' be the largest nonnegative integers such that $\beta - k\alpha$ and $\beta + k'\alpha$ are roots. Then it is known (see Weyl [11(b)]) that $\beta(H_\alpha) = k - k'$ and $\beta + r\alpha$ is a root or zero for an integer r if and only if $-k \leqq r \leqq k'$ Moreover if s_α is the Weyl reflexion corresponding to α, $s_\alpha(\beta + k'\alpha) = \beta - k\alpha$. These facts should be constantly borne in mind during the following discussion.

LEMMA 10. *If* $\gamma, \delta \varepsilon P_+$, $\gamma(H_\delta) \geqq 0$.

For $\gamma + \delta$ is not a root (Lemma 11 of [5(e)]) and obviously it is not zero. Hence $\gamma(H_\delta) \geqq 0$.

LEMMA 11. *Let* α *be any root such that* $H_\alpha \varepsilon \mathfrak{a}_{\mathfrak{l}}$. *Then* α *is compact and* $\gamma_i \pm \alpha$ $(1 \leqq i \leqq s)$ *can never be a root.*

Without loss of generality we may assume that $\alpha > 0$. If α is not compact, it must be totally positive and therefore $\alpha + \gamma_i$ is not a root [5(e), Lemma 11]. Since $H_\alpha \varepsilon \mathfrak{a}_{\mathfrak{l}}$, $\gamma_i(H_\alpha) = 0$ and so it follows that $\gamma_i - \alpha$ also cannot be a root or zero. Hence X_α commutes with X_{γ_i}, $X_{-\gamma_i}$ $1 \leqq i \leqq s$. This however is impossible since $\mathfrak{a}_{\mathfrak{p}}$ is maximal abelian in $\mathfrak{p}$. So α must be compact.

Now consider the sequence $\mathfrak{g} = \mathfrak{g}_1 \supset \mathfrak{g}_2 \supset \cdots$ introduced above. We shall prove that $X_\alpha \varepsilon \mathfrak{g}_r$ for every r. For otherwise choose the least $r \geqq 0$ such that $X_\alpha \not\varepsilon \mathfrak{g}_{r+1}$. Since $\mathfrak{g}_1 = \mathfrak{g}$, $r \geqq 1$ and $X_\alpha \varepsilon \mathfrak{g}_r$. In view of the fact that $X_\alpha \not\varepsilon \mathfrak{g}_{r+1}$, it is clear that either $\gamma_r + \alpha$ or $\gamma_r - \alpha$ is a root. But γ_r is the lowest totally positive root γ such that $X_\gamma \varepsilon \mathfrak{g}_r$. Since $X_\alpha \varepsilon \mathfrak{g}_r$, $X_{-\alpha}$ also lies in $\mathfrak{g}_r$ and so if $\gamma_r - \alpha$ were a root, $X_{\gamma_r - \alpha}$ would also lie in $\mathfrak{g}_r$. Since α is compact and positive, $\gamma_r - \alpha$ is also totally positive and $\gamma_r - \alpha < \gamma_r$. As this contradicts the definition of γ_r, we conclude that $\gamma_r - \alpha$ is not a root and therefore $\gamma_r + \alpha$ is a root. But this implies that $\gamma_r(H_\alpha) < 0$ which, in its turn, contradicts the fact $\gamma_r(H) = 0$ for all $H \varepsilon \mathfrak{a}_{\mathfrak{l}}$. Hence the lemma.

LEMMA 12. *Let* α *be a root. Then for any* i $(1 \leqq i \leqq s)$, $\gamma_i + \alpha$ *and* $\gamma_i - \alpha$ *cannot both be roots.*

We may assume $\alpha > 0$. If α is noncompact $\gamma_i + \alpha$ cannot be a root [5(e), Lemma 11]. So we may assume that α is compact. Suppose then that for some i, $\gamma_i \pm \alpha$ are both roots. Then they are both totally positive and hence from Lemma 10,

$$\gamma_i(H_\delta) \pm \alpha(H_\delta) \geqq 0$$

for every $\delta \varepsilon P_+$ $(j \neq i)$. Then it follows from Lemma 8 that $\gamma_i(H_{\gamma_j}) = 0$ and therefore $\pm \alpha(H_{\gamma_j}) \geqq 0$. This means that $\alpha(H_{\gamma_j}) = 0$. On the other hand $\gamma_i + \alpha$, γ_i, $\gamma_i - \alpha$ are all roots and therefore it follows from Lemma 15 of [5(e)] that $\gamma_i + 2\alpha$ and $\gamma_i - 2\alpha$ are not roots. Hence $\gamma_i(H_\alpha) = 0$. But this implies that $\alpha(H_{\gamma_i}) = 0$ and therefore $\alpha(H_{\gamma_j}) = 0$ $1 \leqq j \leqq s$. This however means that $H_\alpha \varepsilon \mathfrak{a}_{\mathfrak{k}}$ and so we get a contradiction with Lemma 11.

Let λ and μ be two linear functions on $\mathfrak{h}$. We write $\lambda \sim \mu$ if $\lambda - \mu$ vanishes identically on $\nu(\mathfrak{a}_{\mathfrak{p}}) = \sum_{1 \leqq i \leqq s} CH_{\gamma_i}$.

LEMMA 13. *Let α be a positive compact root. Then there are only the following three mutually exclusive possibilities:*

(1) *$H_\alpha \varepsilon \mathfrak{a}_{\mathfrak{k}}$ and therefore $\alpha \sim 0$ and $\gamma_i \pm \alpha$ $(1 \leqq i \leqq s)$ is never a root.*

(2) *There exists a unique index i $(1 \leqq i \leqq s)$ such that $\alpha + \frac{1}{2}\gamma_i \sim 0$.*

(3) *There exists two unique indices i, j $(1 \leqq i < j \leqq s)$ such that $\alpha \sim \frac{1}{2}(\gamma_j - \gamma_i)$.*

Since the first case is covered by Lemma 11, we may assume that $H_\alpha \not\varepsilon \mathfrak{a}_{\mathfrak{k}}$. Then $\alpha(H_{\gamma_i}) \neq 0$ for some i and therefore $X_\alpha \not\varepsilon \mathfrak{g}_{s+1}$. Let i be the least index $(1 \leqq i \leqq s)$ such that $X_\alpha \not\varepsilon \mathfrak{g}_i$. Since α is positive, $\gamma_i + \alpha$ is a root while $\gamma_i - \alpha$ is not (see the proof of Lemma 11). Now suppose $\gamma_j + \epsilon\alpha$ is a root for some j $(1 \leqq j \leqq s, \epsilon = \pm 1)$. If $j \neq i$ we claim $\epsilon = -1$. For otherwise suppose $\gamma_j + \alpha$ is a root. Then it follows from Lemmas 9 and 10 that $\alpha(H_{\gamma_i}) = \gamma_j(H_{\gamma_i}) + \alpha(H_{\gamma_i}) \geqq 0$. On the other hand since $\gamma_i + \alpha$ is a root while $\gamma_i - \alpha$ is not, it is clear that $\gamma_i(H_\alpha) < 0$ and therefore $\alpha(H_{\gamma_i}) < 0$. As this conflicts with our conclusion above, $\epsilon = -1$. So we have two cases. Either (1) $\gamma_j \pm \alpha$ is never a root for $j \neq i$ or (2) $\gamma_j - \alpha$ is a root for some $j \neq i$.

In the first case $\alpha(H_{\gamma_j}) = 0$ for all $j \neq i$. Moreover α, $\alpha + \gamma_i$ are roots while $\alpha - \gamma_i$ is not. Since γ_i and $\alpha + \gamma_i$ are both totally positive $\alpha + 2\gamma_i$ is not a root [5(e), Lemma 11]. Hence $\alpha(H_{\gamma_i}) = -1$. This shows that $\alpha(H_{\gamma_j}) + \frac{1}{2}\gamma_i(H_{\gamma_j}) = 0$ for all j $(1 \leqq j \leqq s)$ and therefore $\alpha + \frac{1}{2}\gamma_i \sim 0$.

Now consider the second case. Let j be the least index such that $\gamma_j - \alpha$ is a root. Then $j \neq i$ and in view of our definition of i, $j > i$. If k

is any index $(1 \leqq k \leqq s)$ other than i, j we claim $\gamma_k \pm \alpha$ cannot be a root. We have already seen this for $\gamma_k + \alpha$. So suppose $\gamma_k - \alpha$ is a root. Then $\gamma_k(H_{\gamma_j}) - \alpha(H_{\gamma_j}) \geqq 0$ (Lemma 10) and therefore $-\alpha(H_{\gamma_j}) \geqq 0$ (Lemma 8). On the other hand α, $\alpha - \gamma_j = -(\gamma_j - \alpha)$ are roots while $\alpha + \gamma_j$ is not. Therefore $\alpha(H_{\gamma_j}) > 0$ giving a contradiction. This proves that $\gamma_k \pm \alpha$ are never roots $(k \neq i, j)$. Moreover as we have seen above, α, $\alpha + \gamma_i$ are roots while $\alpha - \gamma_i$ and $\alpha + 2\gamma_i$ are not and therefore $\alpha(H_{\gamma_i}) = -1$. Similarly since α, $\alpha - \gamma_j$ are roots while $\alpha + \gamma_j$, $\alpha + 2\gamma_j$ are not, $\alpha(H_{\gamma_j}) = 1$. Finally $\alpha(H_{\gamma_k}) = 0$ $(k \neq i, j)$ in view of our result above. Hence it is clear that $\alpha(H_{\gamma_k}) = \frac{1}{2}\gamma_j(H_{\gamma_k}) - \frac{1}{2}\gamma_i(H_{\gamma_k})$ $(1 \leqq k \leqq s)$ and therefore $\alpha \sim \frac{1}{2}(\gamma_j - \gamma_i)$. Moreover since γ_k $(1 \leqq k \leqq s)$ vanish identically on $\mathfrak{a}_{\mathfrak{k}}$, it is clear that their restriction on $\nu(\mathfrak{a}_{\mathfrak{p}})$ are linearly independent. The uniqueness of the indices i and j in the second and third cases of our lemma and the mutual exclusiveness of the three possibilities are therefore obvious.

For any index i $(1 \leqq i \leqq s)$ let C_i denote the set of all compact roots α such that $\alpha + \frac{1}{2}\gamma_i \sim 0$. Similarly let P_i denote the set of all totally positive roots γ for which $\gamma \sim \frac{1}{2}\gamma_i$. If $\alpha \varepsilon C_i$, it follows from Lemma 13 that $-\alpha$ cannot be positive. This shows that C_i consists of positive roots.

LEMMA 14. *$\alpha \rightarrow \gamma_i + \alpha$ $(\alpha \varepsilon C_i)$ is a one-one mapping of C_i onto P_i $(1 \leqq i \leqq s)$.*

For any root β let s_β denote the Weyl reflexion corresponding to β. Now if $\alpha \varepsilon C_i$, $\alpha \sim -\frac{1}{2}\gamma_i$ and therefore $\alpha(H_{\gamma_i}) = -1$. Hence $s_{\gamma_i}\alpha = \alpha - \alpha(H_{\gamma_i})\gamma_i = \alpha + \gamma_i$ and so $\gamma_i + \alpha$ is a root which is obviously in P_i. Conversely if $\gamma \varepsilon P_i$, $\gamma \sim \frac{1}{2}\gamma_i$ and therefore $\gamma(H_{\gamma_i}) = 1$. Hence $s_{\gamma_i}\gamma = \gamma - \gamma_i$ is a root. But as γ and γ_i are both noncompact, $\alpha = \gamma - \gamma_i$ must be compact. Moreover $\alpha + \frac{1}{2}\gamma_i = \gamma - \frac{1}{2}\gamma_i \sim 0$ and therefore $\alpha \varepsilon C_i$. Since it is obvious from its definition that the mapping is one-one, the lemma is proved.

For any given pair of indices i, j $(1 \leqq i < j \leqq s)$, let C_{ij} denote the set of all compact roots α such that $\alpha \sim \frac{1}{2}(\gamma_j - \gamma_i)$. Similarly let P_{ij} denote the set of all $\gamma \varepsilon P_+$ such that $\gamma \sim \frac{1}{2}(\gamma_j + \gamma_i)$. Again we conclude from Lemma 13 that every root in C_{ij} is positive.

LEMMA 15. *$\alpha \rightarrow \gamma_i + \alpha$ $(\alpha \varepsilon C_{ij})$ is a one-one mapping of C_{ij} onto P_{ij}.*

Let $\alpha \varepsilon C_{ij}$. Then $\alpha \sim \frac{1}{2}(\gamma_j - \gamma_i)$ and therefore $\alpha(H_{\gamma_i}) = -1$. Hence $s_{\gamma_i}\alpha = \alpha + \gamma_i$ and so it is clear that $\gamma_i + \alpha \varepsilon P_{ij}$. Conversely if $\gamma \varepsilon P_{ij}$, $\gamma(H_{\gamma_i}) = 1$ and therefore $s_{\gamma_i}\gamma = \gamma - \gamma_i = \alpha$ (say). Then α is compact and $\alpha \sim \frac{1}{2}(\gamma_j - \gamma_i)$.

Let C_0 be the set of all positive roots α such that $\alpha \sim 0$. We know (Lemma 11) that every root in C_0 is compact.

LEMMA 16. *Let P_0 denote the set $(\gamma_1, \gamma_2, \cdots, \gamma_s)$. Then P is the disjoint union of C_0, C_i, C_{ij}, P_0, P_i, P_{ij} $(1 < i < j < s)$.*

Since $\gamma_1, \cdots, \gamma_s$ are linearly independent on $\nu(\mathfrak{a}_\mathfrak{p})$, it is obvious that these sets are all disjoint. Let Q be their union. Then if γ is any positive root, we have to show that $\gamma \varepsilon Q$. If γ is compact, this follows from Lemma 13. So now suppose $\gamma \varepsilon P_+$. Since $\mathfrak{g}_{s+1} \subset \mathfrak{k}$, we can choose an index i $(1 \leqq i \leqq s)$ such that $X_\gamma \varepsilon \mathfrak{g}_i$ but $X_\gamma \not\varepsilon \mathfrak{g}_{i+1}$. Moreover since $\gamma_i \varepsilon P_0 \subset Q$, we may assume that $\gamma \neq \gamma_i$. Then it follows from the definition of γ_i that $\gamma > \gamma_i$. As both γ and γ_i are in P_+, $\gamma + \gamma_i$ is not a root [5(e), Lemma 11]. Therefore since $X_\gamma \not\varepsilon \mathfrak{g}_{i+1}$, $\alpha = \gamma - \gamma_i$ must be a root which is then obviously compact and positive. Therefore we can apply Lemma 13 to α. Since $\gamma = \gamma_i + \alpha$ is a root, $\alpha \not\varepsilon C_0$ (Lemma 11). Hence either $\alpha \sim -\frac{1}{2}\gamma_j$ or $\alpha \sim \frac{1}{2}(\gamma_k - \gamma_j)$ for some j or (k, j) $(1 \leqq j < k \leqq s)$. In the first case $\gamma \sim \gamma_i - \frac{1}{2}\gamma_j$ and so[7] $\gamma(H_{\gamma_j}) = 2\delta_{ij} - 1$. But we know from Lemma 10, that $\gamma(H_{\gamma_j}) \geqq 0$. Therefore $i = j$, $\gamma \sim \frac{1}{2}\gamma_i$ and $\gamma \varepsilon P_i$. In the second case $\gamma \sim \gamma_i + \frac{1}{2}(\gamma_k - \gamma_j)$ and $\gamma(H_{\gamma_j}) = \delta_{ij} - 1$ since $k \neq j$. Therefore again in view of the fact that $\gamma(H_{\gamma_j}) \geqq 0$, we conclude that $i = j$ and hence $\gamma \varepsilon P_{jk}$. This shows that $\gamma \varepsilon Q$ and therefore $Q = P$.

LEMMA 17. *Let α, β be two roots such that $\alpha = \frac{1}{2}(\gamma_j - \gamma_i)$ and $\beta = \frac{1}{2}(\gamma_k - \gamma_j)$ $(1 \leqq i, j, k \leqq s)$. Then they are both compact if $k \neq i$, $\beta + \alpha = \frac{1}{2}(\gamma_k - \gamma_i)$ is a root.*

Consider the scalar product $\langle \alpha, \beta \rangle$ (see [5(e), § 2]). Since $k \neq i$, it follows from Lemma 8 that $\langle \alpha, \beta \rangle = -\frac{1}{4}\langle \gamma_j, \gamma_j \rangle < 0$. Hence $\alpha(H_\beta) < 0$ and therefore $\alpha + \beta$ is a root. The compactness of α and β is an immediate consequence of Lemma 16.

Let us say that $\gamma_i \prec \gamma_j$ $(1 \leqq i, j \leqq s)$ if $\frac{1}{2}(\gamma_j - \gamma_i)$ is a positive root. The above lemma shows that this relation is transitive and therefore it defines a partial order in the set P_0. It is obvious from the definition of γ_i that $\gamma_i < \gamma_j$ if $i < j$. Hence $\gamma_i \prec \gamma_j$ implies $i < j$.

Let r_i and r_{ij} be the number of roots in C_i and C_{ij} $(1 \leqq i < j \leqq s)$ respectively. Then it follows from Lemmas 14 and 15 that these are also the number of roots in P_i and P_{ij} respectively. Put $2\rho_+ = \sum_{\beta \varepsilon P_+} \beta$. Then we have the following result.

---

[7] $\delta_{ij} = 1$ or 0 according as $i = j$ or not.

LEMMA 18. $2\rho_+(H_{\gamma_i}) = 2 + r_i + \sum_{i<j\leqq s} r_{ij} + \sum_{1\leqq j<i} r_{ji}$ $(1 \leqq i \leqq s)$.

Let Q_i be the union of P_i, P_{ij} $(i<j\leqq s)$ and P_{ji} $(1\leqq j<i)$. Put $2\rho_i = \sum_{\gamma\varepsilon Q_i} \gamma$. Then since $\gamma_j(H_{\gamma_i}) = 2\delta_{ji}$, it is obvious that

$$2\rho_i(H_{\gamma_i}) = r_i + \sum_{i<j\leqq s} r_{ij} + \sum_{1\leqq j<i} r_{ji}.$$

Also if γ is a totally positive root which does not lie in Q_i, it follows from Lemma 16 that $\gamma(H_{\gamma_i}) = 0$ unless $\gamma = \gamma_i$. Therefore since $\gamma_i(H_{\gamma_i}) = 2$,

$$2\rho_i(H_{\gamma_i}) = 2 + 2\rho_i(H_{\gamma_i})$$

and this gives the result.

LEMMA 19. *r_{ij} $(1\leqq i<j\leqq s)$ is even if and only if $\frac{1}{2}(\gamma_j - \gamma_i)$ is not a root.*

Let θ' denote the automorphism $\nu\theta\nu^{-1}$ of $\mathfrak{g}$. Since $\theta(\mathfrak{a}) = \mathfrak{a}$ and $\mathfrak{h} = \nu(\mathfrak{a})$, it follows that $\theta'(\mathfrak{h}) = \mathfrak{h}$. Therefore if α is a root the linear function $H \rightarrow \alpha(\theta' H)$ $(H \varepsilon \mathfrak{h})$ is also a root. We denote it by $\theta'\alpha$. It is clear that $\theta' H = -H$ if $H \varepsilon \nu(\mathfrak{a}_{\mathfrak{p}})$ and $\theta' H = H$ if $H \varepsilon \mathfrak{a}_{\mathfrak{k}}$. Hence $\alpha \neq -\theta'\alpha$ unless α vanishes identically on $\mathfrak{a}_{\mathfrak{k}}$. Now suppose $\alpha \varepsilon C_{ij}$ $(1\leqq i<j\leqq s)$. Then $\alpha \sim \frac{1}{2}(\gamma_j - \gamma_i)$ and therefore it is obvious that $\theta'\alpha \sim -\frac{1}{2}(\gamma_j - \gamma_i)$. In view of Lemma 16, this implies that $-\theta'\alpha \varepsilon C_{ij}$. Hence the mapping $\alpha \rightarrow -\theta'\alpha$ defines a permutation of order 2 in the set C_{ij}. Moreover $\alpha \neq -\theta'\alpha$ unless $\alpha = \frac{1}{2}(\gamma_j - \gamma_i)$. Therefore if we pair off α and $-\theta'\alpha$ together, it follows immediately that r_{ij} is odd or even according as $\frac{1}{2}(\gamma_j - \gamma_i)$ is a root or not.

7. Digression on a theorem of Cartan.

Put

$$\mathfrak{p}_+ = \sum_{\beta\varepsilon P_+} CX_\beta \text{ and } \mathfrak{p}_- = \sum_{\beta\varepsilon P_+} CX_{-\beta}.$$

Then $\mathfrak{p}_+$, $\mathfrak{p}_-$ are abelian subalgebras of $\mathfrak{g}$ [5(e), Lemma 11] and $\mathfrak{g}$ is the direct sum of $\mathfrak{k}$, $\mathfrak{p}_+$ and $\mathfrak{p}_-$. Let G_c denote the simply connected complex Lie group with the Lie algebra $\mathfrak{g}$ and let $\mathfrak{P}_c^+$, K_c, $\mathfrak{P}_c^-$ be its analytic subgroups corresponding to $\mathfrak{p}_+$, $\mathfrak{k}$, $\mathfrak{p}_-$ respectively. Also let G_0, K_0 be the real analytic subgroups of G_c corresponding to $\mathfrak{g}_0$, $\mathfrak{k}_0$ respectively. Then $(q, k, p) \rightarrow qkp$ $(q \varepsilon \mathfrak{P}_c^-, k \varepsilon K_c, p \varepsilon \mathfrak{P}_c^+)$ is a one-one regular holomorphic mapping of the complex manifold $\mathfrak{P}_c^- \times K_c \times \mathfrak{P}_c^+$ into G_c and G_0 is contained in $\mathfrak{P}_c^- K_c \mathfrak{P}_c^+$ [5(f), Lemmas 4 and 5].

LEMMA 20. *Let* $X = \sum_{i=1}^{s} t_i(X_{\gamma_i} + X_{-\gamma_i})$ $(t \varepsilon C)$. *Then*

$$\exp X = \exp Y \exp H \exp Z$$

in G_c *where* [6]

$$Y=\sum_{i=1}^{s}(\tanh t_i)X_{-i}, \qquad Z=\sum_{i=1}^{s}(\tanh t_i)X_i \qquad H=\sum_{i=1}^{s}\log(\cosh t_i)H_i$$

provided $\cosh t_i \neq 0$ $1 \leqq i \leqq s$.

This follows immediately from Lemmas 8 and 9.

Now if we put $(X,Y)=-B(\tilde{\theta}(X),Y)$ and $\| X \| = (X,X)^{\frac{1}{2}}$ $(X, Y \varepsilon \mathfrak{g})$, $\mathfrak{g}$ becomes a finite-dimensional Hilbert space. Moreover since $\mathrm{ad}X$ is nilpotent for $X \varepsilon \mathfrak{p}_-$, it is easy to see that $X \rightarrow \exp X$ $(X \varepsilon \mathfrak{p}_-)$ is a one-one regular holomorphic mapping of $\mathfrak{p}_-$ onto $\mathfrak{P}_c^-$. Let $q \rightarrow \log q$ $(q \varepsilon \mathfrak{P}_c^-)$ denote its inverse. For $x \varepsilon G_0$, let $\zeta(x)$ denote the unique element in $\mathfrak{P}_c^-$ such that $x \varepsilon \zeta(x) K_c \mathfrak{P}_c^+$. Then we have the following result.

LEMMA 21. $\| \log \zeta(x) \|$ *remains bounded as* x *varies in* G_0.

Let $\mathfrak{P}_0$ be the set of all elements in G_0 of the form $\exp X$ $(X \varepsilon \mathfrak{p}_0)$. Then it is known that $G_0 = K_0 \mathfrak{P}_0$ (see Cartan [2(b), p. 17], also Mostow [9]). Let $z \rightarrow \mathrm{Ad}(z)$ $(z \varepsilon G_c)$ denote the adjoint representation of G_c. It follows from the definition of $\tilde{\theta}$ that if $k \varepsilon K_0$, $\mathrm{Ad}(k)$ is a unitary operator on $\mathfrak{g}$. Now $\mathfrak{a}_{\mathfrak{p}_0}$ is a maximal abelian subspace of $\mathfrak{p}_0$ (Lemma 8) and therefore $\mathfrak{p}_0 = \bigcup_{k \varepsilon K} \mathrm{Ad}(k)\mathfrak{a}_{\mathfrak{p}_0}$ (see Lemma 33 and also Cartan [2(a), p. 359]). Hence $G_0 = K_0 \mathfrak{A} K_0$ where $\mathfrak{A}$ is the analytic subgroup of G_0 corresponding to $\mathfrak{a}_{\mathfrak{p}_0}$. Moreover since $[\mathfrak{k}, \mathfrak{p}_-] \subset \mathfrak{p}_-$, it is obvious that $\zeta(kxk') = k\zeta(x)k^{-1}$ $(k, k' \varepsilon K_0$, $x \varepsilon G_0)$. Therefore if $x = kak'$ $(k, k \varepsilon K_0 ; a \varepsilon \mathfrak{A})$,

$$\log \zeta(x) = \mathrm{Ad}(k)(\log \zeta(a))$$

and so

$$\| \log \zeta(x) \| = \| \log \zeta(a) \|.$$

Now suppose $a = \exp X$ where $X = \sum_{i=1}^{s} t_i (X_{\gamma_i} + X_{-\gamma_i})$ $(t_i \varepsilon R)$. Then from Lemma 26,

$$\log \zeta(a) = \sum_{i=1}^{s} (\tanh t_i) X_{-\gamma_i}$$

and therefore

$$\| \log \zeta(a) \| \leqq \sum_{i=1}^{s} \| X_{-\gamma_i} \|$$

since $|\tanh t| \leqq 1$ for real t. Thus

$$\| \log \zeta(x) \| \leqq \sum_{i=1}^{s} \| X_{-\gamma_i} \|$$

for all $x \varepsilon G_0$ and so the lemma is proved.

This result has the following significance in relation to the theory of bounded symmetric homogeneous domains of E. Cartan [2(c)]. We know that $G_0 K_c \mathfrak{P}_c^+$ is open in $\mathfrak{P}_c^- K_c \mathfrak{P}_c^+$ and $G_0 \cap (K_c \mathfrak{P}_{c+}) = K_0$ (see [5(f), § 2]). Since $K_c \mathfrak{P}_c^+$ is a group and $\mathfrak{P}_c^- \cap (K_c \mathfrak{P}_c^+) = \{1\}$, we can identify $\mathfrak{P}_c^-$ with the factor space $(\mathfrak{P}_c^- K_c \mathfrak{P}_c^+)/K_c \mathfrak{P}_c^+$. In this way $G_0/K_0 = (G_0 K_c \mathfrak{P}_c^+)/K_c \mathfrak{P}_c^+$ becomes an open submanifold of $\mathfrak{P}_c^-$. The above lemma then shows that this submanifold is equivalent to a *bounded* domain in the complex Euclidean space $\mathfrak{p}_-$. This fact had previously been verified by Cartan [2(c)] by using the classification of all real simple groups and constructing the domain in each case separately.

8. **Transformation of certain integrals.** Let $(-1)^{\frac{1}{2}}$ denote a fixed square-root of -1 in C and put $\mathfrak{u} = \mathfrak{k}_0 + (-1)^{\frac{1}{2}} \mathfrak{p}_0$. Then $\mathfrak{u}$ is a compact real form of $\mathfrak{g}$ (see [5(b), p. 187]). Let $\mathfrak{a}_{\mathfrak{p}_0}$ denote any (real) maximal abelian subspace of $\mathfrak{p}_0$. We denote by $\mathfrak{a}_\mathfrak{p}$ the complexification of $\mathfrak{a}_{\mathfrak{p}_0}$ in $\mathfrak{p}$. Define G_c, G_0, K_0 as in Section 7 and let U, $\mathfrak{A}$, $\mathfrak{A}^*$ and $\mathfrak{A}_c$ be the (real) analytic subgroups of G_c corresponding to $\mathfrak{u}$, $\mathfrak{a}_{\mathfrak{p}_0}$, $(-1)^{\frac{1}{2}} \mathfrak{a}_{\mathfrak{p}_0}$ and $\mathfrak{a}_\mathfrak{p}$ respectively. Then K_0, U and $\mathfrak{A}^*$ are compact (see § 12 and [5(f), § 2]). Put $\mathfrak{q} = [\mathfrak{z}, \mathfrak{a}_\mathfrak{p}]$. It is obvious that $\mathrm{Ad}(a)\mathfrak{q} = \mathfrak{q}$ for $a \in \mathfrak{A}_c$. We put

$$D(a) = \det(\mathrm{Ad}(a) - \mathrm{Ad}(a^{-1}))_\mathfrak{q} \qquad (a \in \mathfrak{A}_c)$$

where $(\mathrm{Ad}(a) - \mathrm{Ad}(a^{-1}))_\mathfrak{q}$ is the restriction of $\mathrm{Ad}(a) - \mathrm{Ad}(a^{-1})$ on $\mathfrak{q}$. Let dx, dk, du, da, da^* denote the Haar measures on G_0, K_0 U, and $\mathfrak{A}^*$ respectively. We assume that

$$\int_{K_0} dk = \int_U du = 1.$$

On the other hand da and da^* are normalized as follows. The metric on $\mathfrak{g}$ (see Section 7) defines a Euclidean metric on the real vector space $\mathfrak{a}_{\mathfrak{p}_0}$ which is given by $\| H \|^2 = B(H, H)$ $(H \in \mathfrak{a}_{\mathfrak{p}_0})$. Let dH denote the element of volume in $\mathfrak{a}_{\mathfrak{p}_0}$ corresponding to this Euclidean metric and put $e(H) = \exp(-1)^{\frac{1}{2}} H$. The mappings $H \to \exp H$ and $H \to e(H)$ $(H \in \mathfrak{a}_{\mathfrak{p}_0})$ define homomorphisms of the additive group $\mathfrak{a}_{\mathfrak{p}_0}$ onto $\mathfrak{A}$ and $\mathfrak{A}^*$ respectively and it is clear that these homomorphisms are local isomorphisms. Hence we can normalize the Haar measures da and da^* in such a way that $da = dH = da^*$ $(a = \exp H$ and $a^* = e(H)$, $H \in \mathfrak{a}_{\mathfrak{p}_0})$.

Let $\mathfrak{A}'$ and $\mathfrak{A}^{*\prime}$ be the sets of those points a in $\mathfrak{A}$ and $\mathfrak{A}^*$ respectively where $D(a) \neq 0$. Then both $\mathfrak{A}'$ and $\mathfrak{A}^{*\prime}$ have only a finite number of connected components (see Section 12). Let w and w^* respectively denote

their number. Moreover let $C_c(G_0)$ be the set of all continuous functions on G_0 which vanish outside a compact set.

LEMMA 22. *Let g be a continuous function on U and $B_0{}^*$ a connected component of $\mathfrak{A}^{*\prime}$. Then*

$$\int_U g(u)\,du \int_{\mathfrak{A}^*} |D(a^*)|^{\frac{1}{2}}\,da^* = w^* \int_{B_0^*} |D(a^*)|^{\frac{1}{2}}\,da^* \int_{K_0\times K_0} g(ka^*k')\,dk\,dk'.$$

Moreover we can normalize the Haar measure dx on G_0 in such a way that

$$\int_{G_0} f(x)dx = w \int_{B_0} |D(a)|^{\frac{1}{2}}\,da \int_{K_0\times K_0} f(kak')\,dk\,dk'$$

for all $f \varepsilon C_c(G_0)$ and every connected component B_0 of $\mathfrak{A}'$. This normalization of dx and the numbers w, w^* and $\int_{\mathfrak{A}^*} |D(a^*)|^{\frac{1}{2}}\,da^*$ are independent of the choice of $\mathfrak{a}_{\mathfrak{p}_0}$.

Although the proof of this lemma is not difficult, due to some technical complications, it is rather long. Hence in order not to interrupt our main argument, we postpone it until Section 12.

Now we assume that $\mathfrak{a}_{\mathfrak{p}_0}$, $\mathfrak{a}_{\mathfrak{p}}$, $\mathfrak{a}_{\mathfrak{k}}$ and $\mathfrak{a}$ are defined as in Section 6 so that $\nu(\mathfrak{a}) = \mathfrak{h}$. Let Σ be the set of all roots of $\mathfrak{g}$ with respect to $\mathfrak{a}$, which do not vanish identically on $\mathfrak{a}_{\mathfrak{p}}$. Then it is obvious that

$$|D(\exp H)| = |\prod_{\alpha\varepsilon\Sigma} (e^{\alpha(H)} - e^{-\alpha(H)})| \qquad (H \varepsilon \mathfrak{a}_{\mathfrak{p}}).$$

Now every linear function λ on $\mathfrak{h}$ defines a linear function λ' on $\mathfrak{a}$ by the rule $\lambda'(H) = \lambda(\nu(H))$ $(H \varepsilon \mathfrak{a})$. Moreover since $\mathfrak{a} \cap \mathfrak{h} = \mathfrak{a}_{\mathfrak{k}}$ and $\nu(H) = H$ for $H \varepsilon \mathfrak{a}_{\mathfrak{k}}$, λ and λ' coincide on $\mathfrak{a} \cap \mathfrak{h}$. Finally since ν is an automorphism of $\mathfrak{g}$, it is obvious that λ' is a root of $\mathfrak{g}$ with respect to $\mathfrak{a}$ if and only if λ is a root with respect to $\mathfrak{h}$. Hence if we identify linear functions on $\mathfrak{h}$ with those on $\mathfrak{a}$ under the mapping $\lambda \to \lambda'$, the two sets of roots coincide. Then Σ is exactly the set of those roots α for which $H_\alpha \not\in \mathfrak{a}_{\mathfrak{k}}$ (in the notation of Section 6). Let Q be the set of those roots in P which are not identically zero on $\nu(\mathfrak{a}_{\mathfrak{p}})$. Then it follows from Lemma 16 that Q is the disjoint union of C_i, C_{ij}, P_0, P_i, P_{ij} $(1 \leqq i < j \leqq s)$. Moreover it is obvious that

$$|D(\exp H)|^{\frac{1}{2}} = |\prod_{\alpha\varepsilon Q} (e^{\alpha(H)} - e^{-\alpha(H)})| \qquad (H \varepsilon \mathfrak{a}_{\mathfrak{p}}).$$

Now put

$$H = \sum_{i=1}^{s} t_i(X_{\gamma_i} + X_{-\gamma_i}) \qquad (t_i \varepsilon C).$$

Then

$$\nu(H)=\sum_{i=1}^{s} t_i H_{\gamma_i} \text{ and therefore}$$

$$\alpha(H)=\alpha(\nu(H))=\sum_{i=1}^{s} t_i \alpha(H_{\gamma_i}) \qquad (\alpha \in Q).$$

Since $\gamma_i(H_{\gamma_j})=2\delta_{ij}$ $(1 \leqq i, j \leqq s)$ it is obvious (see Section 6) that

$$\begin{aligned} \alpha(H)&=-t_i && \text{if } \alpha \in C_i, \\ \alpha(H)&=t_j-t_i && \text{if } \alpha \in C_{ij}, \\ \alpha(H)&=t_i && \text{if } \alpha \in P_i, \\ \alpha(H)&=t_j+t_i && \text{if } \alpha \in P_{ij} \qquad (1 \leqq i<j \leqq s). \end{aligned}$$

Put

$$Q_1=\bigcup_{1 \leqq i \leqq s}(C_i \cup P_i), \qquad Q_2=\bigcup_{1 \leqq i<j \leqq s}(C_{ij} \cup P_{ij}),$$

Then in the notation of Section 6,

$$\Big|\prod_{\alpha \in Q_1}(e^{\alpha(H)}-e^{-\alpha(H)})\Big|=\prod_{1 \leqq i \leqq s} 2^{2r_i}|\sinh t_i|^{2r_i},$$

$$\Big|\prod_{\alpha \in P_0}(e^{\alpha(H)}-e^{-\alpha(H)})\Big|=\prod_{1 \leqq i \leqq s}|4 \sinh t_i \cosh t_i|.$$

Moreover since

$$\sinh(t_j-t_i)\sinh(t_j+t_i)=(\cosh t_j)^2-(\cosh t_i)^2,$$

it follows that

$$\Big|\prod_{\alpha \in Q_2}(e^{\alpha(H)}-e^{-\alpha(H)})\Big|=\prod_{1 \leqq i<j \leqq s} 2^{2r_{ij}}|(\cosh t_j)^2-(\cosh t_i)^2|^{r_{ij}}.$$

Hence

$$\begin{aligned} |D(\exp H)|^{\frac{1}{2}}=&\prod_{1 \leqq i \leqq s} 2^{2(r_i+1)}|\sinh t_i|^{2r_i+1}|\cosh t_i| \\ &\times \prod_{1 \leqq i<j \leqq s} 2^{2r_{ij}}|(\cosh t_j)^2-(\cosh t_i)^2|^{r_{ij}}. \end{aligned}$$

Now suppose $\cosh t_i \neq 0$ $1 \leqq i \leqq s$ and put[6] $t_i'=\log(\cosh t_i)$. Let

$$H'=\sum_{i=1}^{s} t_i' H_{\gamma_i}$$

and consider the expression

$$\Delta(2H')=\prod_{\alpha \in C'}(e^{\alpha(H')}-e^{-\alpha(H')})$$

where C' is the set of all positive compact roots which do not vanish identically on $\nu(\mathfrak{a}_{\mathfrak{p}})$. It follows from Lemma 16 that C' is the disjoint union of C_i and C_{ij} $(1 \leqq i<j \leqq s)$. Therefore it is clear that

$$\Delta(2H') = \prod_i \{\cosh t_i - (1/\cosh t_i)\} \prod_{i<j} \{(\cosh t_j/\cosh t_i) - (\cosh t_i/\cosh t_j)\}^{r_{ij}}$$

$$= \prod_i \{(\sinh t_i)^{2r_i}/(\cosh t_i)^{r_i}\} \prod_{i<j} \{(\cosh t_j)^2 - (\cosh t_i)^2\}^{r_{ij}}/(\cosh t_i \cosh t_j)^{r_{ij}}.$$

Hence

$$| D(\exp H)|^{\frac{1}{2}} = | \Delta(2H')| \{\prod_i 2^{2r_i+2} | \cosh t_i |^{r_i+1} | \sinh t_i |\}$$
$$\times \prod_{i<j} | 4 \cosh t_i \cosh t_j |^{r_{ij}}.$$

But we know from Lemma 18 that

$$r_i + \sum_{i<j\leqq s} r_{ij} + \sum_{1\leqq j<i} r_{ji} = 2\rho_+(H_i) - 2$$

Therefore

$$\{\prod_i 2^{2r_i+2} | \cosh t_i |^{r_i+1}\} \prod_{i<j} | 4 \cosh t_i \cosh t_j |^{r_{ij}}$$
$$= \prod_i 2^{2\rho_+(H_{\gamma_i})} | \cosh t_i |^{2\rho_+(H_{\gamma_i})-1} = 2^p \prod_i | \cosh t_i |^{2\rho_+(H_{\gamma_i})-1}$$

where $p = \sum_{1\leqq i\leqq s} 2\rho_+(H_{\gamma_i})$. Thus we have the following result.

$$| D(\exp H)|^{\frac{1}{2}} = 2^p | \Delta(2H')| \prod_i \{| \cosh t_i |^{2\rho_+(H_{\gamma_i})-1} | \sinh t_i |\}$$

We now introduce the partial order in the set $P_0 = (\gamma_1, \cdots, \gamma_s)$ as described at the end of Section 6. Let $\beta_1, \cdots, \beta_r$ be all the (distinct) minimal elements in P_0 under this order. For any i $(1 \leqq i \leqq r)$ consider the set σ_i of all $\gamma \varepsilon P_0$ such that $\frac{1}{2}(\gamma - \beta_i)$ is a root. Then γ, γ' are two distinct elements of σ_i, it follows from Lemma 17 that either $\gamma \prec \gamma'$ or $\gamma' \prec \gamma$. This shows that σ_i is simply ordered. Therefore we may write it in the form

$$\beta_i = \beta_{i1} \prec \beta_{i2} \cdots \prec \beta_{is_i}.$$

Moreover if $i \neq j$, σ_i, σ_j are disjoint $(1 \leqq i, j \leqq r)$. For otherwise suppose $\gamma \varepsilon \sigma_i \cap \sigma_j$. Then $\frac{1}{2}(\gamma - \beta_i)$ and $\frac{1}{2}(\gamma - \beta_j)$ are both roots and therefore again from Lemma 17, $\frac{1}{2}(\beta_i - \beta_j)$ is a root. This however is impossible since both β_i and β_j are minimal in P_0. Thus P_0 is the disjoint union of $\sigma_1, \cdots, \sigma_r$. We put $t_{ij} = t_k$ if $\beta_{ij} = \gamma_k$ $(1 \leqq i \leqq r, 1 \leqq j \leqq s_i, 1 \leqq k \leqq s)$. Let $\mathfrak{b}$ denote the subset of $\mathfrak{a}_{\mathfrak{p}_0}$ consisting of all elements $H = t_1 H_{\gamma_1} + \cdots + t_s H_{\gamma_s}$ $(t_j \varepsilon R)$ such that

$$0 < t_{i1} < t_{i2} < \cdots < t_{is_i} \qquad (1 \leqq i \leqq r).$$

Similarly let $\mathfrak{b}^*$ denote the subset of $\mathfrak{b}$ consisting of those H which satisfy the additional condition $t_{is_i} < \frac{\pi}{2}$ $(1 \leqq i \leqq r)$. Then we have the following result.

LEMMA 23. *Put* $e(X) = \exp(-1)^{\frac{1}{2}}X$ $(X \varepsilon \mathfrak{g})$ *and*

$$\Delta(H) = \prod_{\alpha \varepsilon C'} (e^{\frac{1}{2}\alpha(H)} - e^{-\frac{1}{2}\alpha(H)}) \qquad (H \varepsilon \mathfrak{h}).$$

For any $H = \sum_i t_i(X_{\gamma_i} + X_{-\gamma_i})$ *in* $\mathfrak{b}$ *let* H' *denote the element*[6]

$$H' = \log(\cosh t_1) H_{\gamma_1} + \cdots + \log(\cosh t_s) H_{\gamma_s}$$

in $\mathfrak{h}$. *Similarly for any* $H = \sum_i t_i(X_{\gamma_i} + X_{-\gamma_i})$ *in* $\mathfrak{b}^*$, *let* H^* *denote the element*

$$H^* = \log(\cos t_1) H_{\gamma_1} + \cdots + \log(\cos t_s) H_{\gamma_s}.$$

Then

$$|D(\exp H)|^{\frac{1}{2}} = 2^p \Delta(2H') \prod_i (\cosh t_i)^{2\rho_+(H\gamma_i)-1} \prod_i \sinh t_i \qquad (H \varepsilon \mathfrak{b})$$

and

$$|D(e(H))|^{\frac{1}{2}} = (-1)^p 2^p \Delta(2H^*) \prod_i (\cos t_i)^{2\rho_+(H\gamma_i)-1} \prod_i \sin t_i \qquad (H \varepsilon \mathfrak{b}^*).$$

If $H \varepsilon \mathfrak{b}$, $t_i > 0$ $(1 \leqq i \leqq s)$ and therefore $\sinh t_i > 0$. Hence in view of our earlier result, in this case it is enough to prove that $\Delta(2H')$ is real and positive. But we have already seen that

$$\Delta(2H') = \prod_i \{(\sinh t_i)^{2r_i}/(\cosh t_i)^{r_i}\} \times \prod_{i<j} \{(\cosh t_j)^2 - (\cosh t_i)^2\}^{r_{ij}}/(\cosh t_i \cosh t_j)^{r_{ij}}.$$

Since $\cosh t$ is a positive increasing function of t for $t > 0$, $\Delta(2H')$ has the same sign as

$$\eta = \prod_{1 \leqq i < j \leqq s} (t_j - t_i)^{r_{ij}}.$$

Now if $\frac{1}{2}(\gamma_j - \gamma_i)$ is not a root, we know from Lemma 19 that r_{ij} is even. On the other hand if $\frac{1}{2}(\gamma_j - \gamma_i)$ is a root, $\gamma_i \prec \gamma_j$ and therefore, in view of the definition of $\mathfrak{b}$, $t_j - t_i > 0$. This shows that $\Delta(2H') \geqq 0$ and so the first assertion of the lemma follows.

Now we come to the second case when $H \varepsilon \mathfrak{b}^*$. Since $\cosh((-1)^{\frac{1}{2}}t) = \cos t$ and $\sinh((-1)^{\frac{1}{2}}t) = (-1)^{\frac{1}{2}} \sin t$ $(t \varepsilon R)$, it follows that

$$\Delta(2H^*) = \prod_i (-1)^{r_i}\{(\sin t_i)^{2r_i}/(\cos t_i)^{r_i}\} \times \prod_{i<j} \{(\cos t_j)^2 - (\cos t_i)^2\}^{r_{ij}}/(\cos t_i \cos t_j)^{r_{ij}}$$

and therefore $\Delta(2H^*)$ is real. Moreover since $\cos t$ is a positive decreasing function of t in the interval $0 < t < \frac{\pi}{2}$ it follows that $\Delta(2H^*)$ has the same sign as

$$\prod_i (-1)^{r_i} \prod_{i<j} (t_i - t_j)^{r_{ij}} = (-1)^q \eta$$

where $q=\sum_i r_i+\sum_{i<j} r_{ij}$. We have seen above that $\eta\geqq 0$ on $\mathfrak{b}$ and therefore also on $\mathfrak{b}^*$ and from Lemma 18,

$$q=\sum_{i=1}^{s}\{2\rho_+(H_{\gamma_i})-2\}=p-2s.$$

Hence $(-1)^q=(-1)^p$. On the other hand $\cos t$ and $\sin t$ are both positive on the interval $0<t<\frac{1}{2}$ and so the second statement of the lemma is an immediate consequence of our earlier expression for $|D|^{\frac{1}{2}}$.

Let B and B^* be the images in $\mathfrak{A}$ and $\mathfrak{A}^*$ of $\mathfrak{b}$ and $\mathfrak{b}^*$ under the mappings $H\to\exp H$ and $H\to e(H)$ respectively.

LEMMA 24. *$B\cap\mathfrak{A}'$ is both open and closed in $\mathfrak{A}'$. Similarly $B^*\cap\mathfrak{A}^{*\prime}$ is both open and closed in $\mathfrak{A}^{*\prime}$.*

Since $\mathfrak{b}$ and $\mathfrak{b}^*$ are obviously open in $\mathfrak{a}_{\mathfrak{p}_0}$ and since the mappings $H\to\exp H$ and $H\to e(H)$ are regular on $\mathfrak{a}_{\mathfrak{p}_0}$, it follows that B and B^* are open in $\mathfrak{A}$ and $\mathfrak{A}^*$ respectively. Let $\bar{\mathfrak{b}}$ and $\bar{\mathfrak{b}}^*$ respectively denote the closures of $\mathfrak{b}$ and $\mathfrak{b}^*$ in the real Euclidean space $\mathfrak{a}_{\mathfrak{p}_0}$. Then $\bar{\mathfrak{b}}^*$ is compact. Since $H\to\exp H$ is a topological mapping of $\mathfrak{a}_{\mathfrak{p}_0}$ onto $\mathfrak{A}$ (see Section 12), the image $\exp\bar{\mathfrak{b}}$ of $\bar{\mathfrak{b}}$ under this mapping is closed in $\mathfrak{A}$. Also $e(\bar{\mathfrak{b}}^*)$ is compact and therefore closed in $\mathfrak{A}^*$. Hence it is enough to prove that

$$(\exp\bar{\mathfrak{b}})\cap\mathfrak{A}'=B\cap\mathfrak{A}',\qquad e(\bar{\mathfrak{b}}^*)\cap\mathfrak{A}^{*\prime}=B^*\cap\mathfrak{A}^{*\prime}.$$

Now let $H=\sum_{i=1}^{s}t_i(X_{\gamma_i}+X_{-\gamma_i})$ $(t_i\,\varepsilon\,R)$ be a point in $\bar{\mathfrak{b}}$. Then $0\leqq t_i$ and if $\gamma_i\prec\gamma_j$, $t_i\leqq t_j$. On the other hand we have seen that

$$|D(\exp H)|^{\frac{1}{2}}=\prod_i 2^{2(r_i+1)}|\sinh t_i|^{2r_i+1}|\cosh t_i|$$
$$\times\prod_{i<j}2^{2r_{ij}}|(\cosh t_j)^2-(\cosh t_i)^2|^{r_{ij}}.$$

Hence if $D(\exp H)\neq 0$, $t_i\neq 0$ and $t_i\neq t_j$ if $r_{ij}>0$ $(1\leqq i<j\leqq s)$. This proves that in this case $H\,\varepsilon\,\mathfrak{b}$ and therefore $(\exp\bar{\mathfrak{b}})\cap\mathfrak{A}'=B\cap\mathfrak{A}'$. Now suppose H lies in $\bar{\mathfrak{b}}^*$. Then we have the additional conditions $0\leqq t_i\leqq\frac{\pi}{2}$ $(i=1,\cdots,s)$. Moreover again we know that

$$|D(e(H))|^{\frac{1}{2}}=\prod_i 2^{2(r_i+1)}|\sin t_i|^{2r_i+1}|\cos t_i|$$
$$\times\prod_{i<j}2^{2r_{ij}}|(\cos t_j)^2-(\cos t_i)^2|^{r_{ij}}$$

and therefore if $D(e(H))\neq 0$, $t_i\neq 0$, $\frac{\pi}{2}$ and $t_i\neq t_j$ whenever $r_{ij}>0$

$(1 \leqq i < j \leqq s)$. This shows that $H \varepsilon \mathfrak{b}^*$ in this case and therefore

$$e(\bar{\mathfrak{b}}) \cap \mathfrak{A}^{*\prime} = B^* \cap \mathfrak{A}^{*\prime}.$$

COROLLARY. *Every connected component of $B \cap \mathfrak{A}'$ or $B^* \cap \mathfrak{A}^{*\prime}$ is also a connected component of $\mathfrak{A}'$ or $\mathfrak{A}^{*\prime}$ respectively. The number of connected components of $B \cap \mathfrak{A}'$ is the same as that of $B^* \cap \mathfrak{A}^{*\prime}$.*

The first statement is obvious from the above lemma. Now first we claim that the mapping $H \to e(H)$ is univalent on $\mathfrak{b}^*$. For suppose $e(H_1) = e(H_2)$ $(H_1, H_2 \varepsilon \mathfrak{b}^*)$. Then it is obvious that if α is any root in Σ, $\alpha(H_1) - \alpha(H_2)$ must be an integral multiple of 2π. Hence in particular $\gamma_j(\nu(H_1)) - \gamma_j(\nu(H_2)) = 2n_j\pi$ $(1 \leqq j \leqq s)$ where n_j is an integer. But since $H_1, H_2 \varepsilon \mathfrak{b}^*$, $0 < \gamma_j(\nu(H_i)) < \pi$ $i = 1, 2$ and therefore $n_j = 0$ $(1 \leqq j \leqq s)$. This however implies that $H_1 = H_2$. Since we already know that the mapping $H \to e(H)$ of $\mathfrak{b}^*$ onto B^* is open and continuous it follows that it is topological.

Now let σ_α $(\alpha \varepsilon \Sigma)$ denote the hyperplane in $\mathfrak{a}_{\mathfrak{p}_0}$ consisting of all points H such that $\alpha(H) = 0$. Let $\mathfrak{a}'_{\mathfrak{p}_0}$ denote the complement of $\bigcup_{\alpha \varepsilon \Sigma} \sigma_\alpha$ in $\mathfrak{a}_{\mathfrak{p}_0}$. Then it is obvious that $D(\exp H) \neq 0$ $(H \varepsilon \mathfrak{a}_{\mathfrak{p}_0})$ if and only if $H \varepsilon \mathfrak{a}'_{\mathfrak{p}_0}$. Since the exponential mapping of $\mathfrak{a}_{\mathfrak{p}_0}$ onto $\mathfrak{A}$ is topological, $B \cap \mathfrak{A}'$ and $\mathfrak{b} \cap \mathfrak{a}'_{\mathfrak{p}_0}$ have the same number of components. Let $\mathfrak{b}_0$ be a connected component of $\mathfrak{b} \cap \mathfrak{a}'_{\mathfrak{p}_0}$. Then it is obvious that every root $\alpha \varepsilon \Sigma$ must keep constant sign on $\mathfrak{b}_0$. Therefore since $\mathfrak{b}$ is obviously a convex set, the same holds for $\mathfrak{b}_0$. Moreover if $H \varepsilon \mathfrak{b}_0$, it is obvious that the half-line consisting of all points tH $(t > 0)$ lies entirely in $\mathfrak{b}_0$. But if t is sufficiently small and positive $tH \varepsilon \mathfrak{b}^*$. This implies that $\mathfrak{b}_0$ contains some connected component of $\mathfrak{b}^* \cap \mathfrak{a}'_{\mathfrak{p}_0}$. Also if H_1, H_2 are two points in $\mathfrak{b}^*$ which both lie in $\mathfrak{b}_0$, the straight line-segment J joining them is also contained in $\mathfrak{b}_0$. Hence $J \subset \mathfrak{a}'_{\mathfrak{p}_0}$. But since $\mathfrak{b}^*$ is obviously convex, $J \subset \mathfrak{b}^* \cap \mathfrak{a}'_{\mathfrak{p}_0}$. This shows that $\mathfrak{b}_0$ cannot contain two distinct connected components of $\mathfrak{b}^* \cap \mathfrak{a}'_{\mathfrak{p}_0}$ and therefore every component of $\mathfrak{b} \cap \mathfrak{a}'_{\mathfrak{p}_0}$ contains exactly one component of $\mathfrak{b}^* \cap \mathfrak{a}'_{\mathfrak{p}_0}$. Conversely it is obvious that every component of $\mathfrak{b}^* \cap \mathfrak{a}'_{\mathfrak{p}_0}$ is contained in exactly one component of $\mathfrak{b} \cap \mathfrak{a}'_{\mathfrak{p}_0}$. Hence these two sets have the same number of components.

On the other hand if $H = \sum_{i=1}^{s} t_i(X_{\gamma_i} + X_{-\gamma_i})$ lies in $\mathfrak{b}^*$,

$$| D(e(H)) |^{\frac{1}{2}} = \prod_i 2^{2(r_i+1)} | \sin t_i |^{2r_i+1} | \cos t_i |$$
$$\times \prod_{i<j} 2^{2r_{ij}} | (\cos t_j)^2 - (\cos t_i)^2 |^{r_{ij}}.$$

Since $0 < t_i < \frac{\pi}{2}$, $D(e(H)) = 0$ if and only if $t_i = t_j$ for some pair of indices i, j $(i < j)$ with $r_{ij} > 0$. But clearly this happens if and only if $\alpha(H) = 0$ for some $\alpha \varepsilon \Sigma$. Hence $D(e(H)) \neq 0$ if and only if $H \varepsilon \mathfrak{b}^* \cap \mathfrak{a}'_{\mathfrak{p}_0}$ and so the mapping $H \rightarrow e(H)$ maps $\mathfrak{b}^* \cap \mathfrak{a}'_{\mathfrak{p}_0}$ topologically onto $B^* \cap \mathfrak{A}^{*\prime}$. Similarly $H \rightarrow \exp H$ $(H \varepsilon \mathfrak{b} \cap \mathfrak{a}'_{\mathfrak{p}_0})$ maps $\mathfrak{b} \cap \mathfrak{a}'_{\mathfrak{p}_0}$ topologically onto $B \cap \mathfrak{A}'$. Therefore $B \cap \mathfrak{A}'$, $\mathfrak{b} \cap \mathfrak{a}'_{\mathfrak{p}_0}$, $\mathfrak{b}^* \cap \mathfrak{a}'_{\mathfrak{p}_0}$, $B^* \cap \mathfrak{A}^{*\prime}$ all have the same number of components.

9. Application to representations of G. Let G be the simply connected covering group of G_0 and let K be the analytic subgroup of G corresponding to $\mathfrak{k}_0$. Define the complex manifold W containing G as in [5(f), § 3] and the mapping Γ of G into the center $\mathfrak{c}$ of $\mathfrak{k}$ as in [5(f), § 6]. Then $\Gamma(xu) = \Gamma(ux) = \Gamma(u) + \Gamma(x)$ $(u \varepsilon K, x \varepsilon G)$ [5(f), Lemma 13] and if λ is a real linear function on $\mathfrak{h}$ (see [5(e), § 2]) $\lambda(\Gamma(u))$ is pure imaginary [5(f), Lemma 23]. Let Λ be a real linear function on $\mathfrak{h}$ such that $\Lambda(H_\alpha)$ is a non-negative integer for every positive compact root α. Consider a fundamental system $(\alpha_1, \cdots, \alpha_l)$ of positive roots and suppose that $(\alpha_1, \cdots, \alpha_m)$ are all the totally positive roots in this system. Let Λ_0 denote the linear function on $\mathfrak{h}$ given by $\Lambda_0(H_{\alpha_i}) = 0$ $1 \leqq i \leqq m$ and $\Lambda_0(H_{\alpha_i}) = \Lambda(H_{\alpha_i})$ $m < i \leqq l$. Then we can define (see [5(f), § 6]) an irreducible complex representation σ of G_c on a finite-dimensional Hilbert space V such that σ is unitary on U and its highest weight is Λ_0. Let ϕ_0 be a unit vector in V belonging to the highest weight Λ_0 and put $\lambda = \Lambda - \Lambda_0$. Then if $x \rightarrow \bar{x}$ $(x \varepsilon G)$ denotes the natural mapping of G onto G_0, we consider the function

$$\psi_\Lambda(x) = (\phi_0, \sigma(\bar{x})\phi_0) e^{\lambda(\Gamma(x))} \qquad (x \varepsilon G)$$

on G. Since the center of G lies in K and Λ is real, $|\psi_\Lambda(x)|$ depends only on $\bar{x}$. We propose to consider the integral (cf. [5(f), § 9])

$$\int_{G_0} |\psi_\Lambda(x)|^2 \, d\bar{x}$$

where $d\bar{x}$ is the Haar measure on G_0. For any $a \varepsilon \mathfrak{A}$ let $\log a$ denote the unique element $H \varepsilon \mathfrak{a}_{\mathfrak{p}_0}$ such that $a = \exp H$. We put $\Gamma(a) = \Gamma(\exp' \log a)$ where $X \rightarrow \exp' X$ $(X \varepsilon \mathfrak{g}_0)$ is the exponential mapping of $\mathfrak{g}_0$ into G. Then if we normalize the various Haar measures in accordance with Lemma 22 and take into account the fact that $\lambda(\Gamma(a))$ is real for $a \varepsilon \mathfrak{A}$ [5(f), Lemma 23], it follows that

$$\int_{G_0} |\psi_\Lambda(x)|^2 \, d\bar{x} = wN^{-1} \int_B e^{2\lambda(\Gamma(a))} \, |D(a)|^{\frac{1}{2}} \, da \int_{K_0 \times K_0} (\phi_0, \sigma(kak')\phi_0)^2 \, dk dk'$$

where B is defined[8] as in Lemma 24 and N is the number of connected components of $B \cap \mathfrak{A}'$. Thus we are led to the integral

$$\int_{K_0 \times K_0} |(\phi^*, \sigma(kak')\phi_0)|^2 \, dk dk' \qquad (a \in B).$$

Let $H = \sum_i t_i(X_{\gamma_i} + X_{-\gamma_i}) \in \mathfrak{h}$ $(t_i \in R)$ and let $a = \exp H$. Then we know from Lemma 20 that

$$a = \zeta h(a) \xi$$

where $\zeta \in \mathfrak{P}_c^-$, $\xi \in \mathfrak{P}_c^+$ (in the notation of Section 7) and

$$h(a) = \exp\left(\sum_{i=1}^{s} \log(\cosh t_i) H_i \right).$$

If we denote the corresponding representation of $\mathfrak{g}$ also by σ, it is obvious that $\sigma(\mathfrak{p}_+)\phi_0 = 0$ and therefore $\sigma(p)\phi_0 = \phi_0$ for $p \in \mathfrak{P}_c^+$. Since $k\mathfrak{P}_c^- k^{-1} = \mathfrak{P}_c^-$, $k\mathfrak{P}_c^+ k^{-1} = \mathfrak{P}_c^+ (k \in K_c)$ and $\bar{\theta}(\mathfrak{P}_c^-) = \mathfrak{P}_c^+$, it follows that

$$(\phi_0, \sigma(kak')\phi_0) = (\phi_0, \sigma(kh(a)k')\phi_0).$$

Let A_c be the complex analytic subgroup of G_c corresponding to $\mathfrak{h}$. Then $h(a) \in A_c \subset K_c$ and

$$\int_{K_0 \times K_0} |(\phi_0, \sigma(kak')\phi_0)|^2 \, dk dk' = \int_{K_0 \times K_0} |(\phi_0, \sigma(kh(a)k')\phi_0)|^2 \, dk dk'.$$

On the other hand $\sigma(H)\phi_0 = \Lambda_0(H)\phi_0$ $(H \in \mathfrak{h})$ and $\sigma(X_\alpha)\phi_0 = 0$ for any positive root α. Since this holds in particular for every positive compact root, it follows from Lemma 2 of [5(e)] (applied to $\mathfrak{k}$) that the subspace V_0 of V spanned by $\sigma(k)\phi_0$ $(k \in K_c)$ is irreducible under K_c. Let σ_0 denote the corresponding representation of K_c (and $\mathfrak{k}$) on V_0. Then obviously Λ_0 is the highest weight of σ_0. Now if we make use of the Schur orthogonality relations for the irreducible representation of the compact group K_0 on V_0, we find easily that

$$\begin{aligned}\int_{K_0 \times K_0} |(\phi_0, \sigma(kyk')\phi_0)|^2 \, dk dk' &= (\dim V_0)^{-1} \int_{K_0} |\sigma(yk)\phi_0|^2 \, dk \\ &= (\dim V_0)^{-2} \operatorname{Sp} \sigma_0(\bar{\theta}(y) y)\end{aligned}$$

if $y \in K_c$. Hence

$$\int_{K_0 \times K_0} |(\phi_0, \sigma(kak')\phi_0)|^2 \, dk dk' = (\dim V_0)^{-2} \chi_{\Lambda_0}(h(a))^2)$$

---

[8] Here we have to make use of the obvious fact that the complement of $\mathfrak{A}'$ in $\mathfrak{A}$ is of measure zero with respect to da.

where χ_{Λ_0} is the character of σ_0. But if $P_{\mathfrak{k}}$ is the set of all compact positive roots and $\rho_{\mathfrak{k}} = \frac{1}{2} \sum_{\alpha \varepsilon P_{\mathfrak{k}}} \alpha$, we know (see Weyl [11(a)]) that

$$\dim V_0 = \prod_{\alpha \varepsilon P_{\mathfrak{k}}} \{\Lambda_0(H_\alpha) + \rho_{\mathfrak{k}}(H_\alpha)/\rho_{\mathfrak{k}}(H_\alpha)\}$$
$$= \prod_{\alpha \varepsilon P_{\mathfrak{k}}} \{\Lambda(H_\alpha) + \rho_{\mathfrak{k}}(H_\alpha)/\rho_{\mathfrak{k}}(H_\alpha)\}$$

since $\lambda(H_\alpha) = 0$ $(\alpha \varepsilon P_{\mathfrak{k}})$ from Lemma 13 of [5(e)].

We now claim that

$$\lambda(\Gamma(a)) = \sum_{i=1}^{s} \log(\cosh t_i)\lambda(H_{\gamma_i})$$

where (in accordance with our convention) $\log(\cosh t_i)$ is real $(1 \leqq i \leqq s)$. Since both sides are linear in λ. it is enough to prove this under the assumption that $\lambda(H_{\alpha_i})$ $1 \leqq i \leqq m$ are all non-negative integers and $\lambda(H_{\alpha_i}) = 0$ $m < i \leqq l$ (see the proof of Lemma 23 of [5(f)]). But then we can find an irreducible complex representation σ' of G_c on a finite-dimensional Hilbert space V' with the highest weight λ and assume that σ' is unitary on U (see [5(f), § 6]). Let ϕ' be a unit vector in V' belonging to the weight λ. Then, as we have seen during the proof of Lemma 23 of [5(f)],

$$(\phi', \sigma'(\bar{x})\phi') = e^{\lambda(\Gamma(a))} \qquad (x \varepsilon G)$$

and therefore by the argument which we have already used above,

$$e^{\lambda(\Gamma(a))} = (\phi', \sigma'(a)\phi') = (\phi', \sigma'(h(a))\phi')$$
$$= \exp\Big(\sum_{i=1}^{s} (\log\cosh t_i)\lambda(H_{\gamma_i})\Big)$$

since ϕ' belongs to the weight λ. Our assertion now follows from the fact that both $\lambda(\Gamma(a))$ and $\sum_i (\log\cosh t_i)\lambda(H_{\gamma_i})$ are real.

If b and μ are two real numbers and b is positive we define b^μ in the usual way by $b^\mu = \exp(\mu \log b)$. Then the above result shows that

$$e^{\lambda(\Gamma(a))} = \prod_{i=1}^{s} (\cosh t_i)^{2\lambda(H_{\gamma_i})}$$

For any point $H = \sum_{i=1}^{s} t_i(X_{\gamma_i} + X_{-\gamma_i})$ $(t_i \varepsilon R)$ in $\mathfrak{a}_{\mathfrak{p}_0}$, we regard $(t_1, \cdots, t_s)$ as the coordinates of H and denote by dt the measure $dt_1 dt_2 \cdots dt_s$ on $\mathfrak{a}_{\mathfrak{p}_0}$. It is obvious that $dH = c\,dt$ where c is a positive constant and dH is the element of volume corresponding to the Euclidean metric on $\mathfrak{a}_{\mathfrak{p}_0}$ (see Section 8). Hence it follows from the above remarks that

$$\int_{G_0} |\psi_\Lambda(x)|^2 d\bar{x} = wN^{-1}(\dim V_0)^{-1}\int_B \chi_{\Lambda_0}((h(a))^2)e^{2\lambda(\Gamma(a))}|D(a)|^{\frac{1}{2}}da$$

$$= wcN^{-1}(\dim V_0)^{-2}\int_{\mathfrak{b}} \chi_{\Lambda_0}(\exp 2H')\prod_{i=1}^{s}(\cosh t_i)^{2\lambda(H\gamma_i)}|D(\exp H)^{\frac{1}{2}}dt$$

where $H = \sum_{i=1}^{s} t_i(X_{\gamma_i} + X_{-\gamma_i})$ and

$$H' = \sum_{i=1}^{s}(\log\cosh t_i)H_{\gamma_i} \qquad (t_i \varepsilon R).$$

Let $\mathfrak{w}_\mathfrak{k}$ be the subgroup of the Weyl group (of $\mathfrak{g}$ with respect to $\mathfrak{h}$) generated by the Weyl reflexions s_α corresponding to $\alpha \varepsilon P_\mathfrak{k}$. For any $s \varepsilon \mathfrak{w}_\mathfrak{k}$ we define $\epsilon(s) = 1$ or -1 according as the permutation $\alpha \to s\alpha$ of the set of all compact roots α, is even or odd. As in Lemma 16, let C_0 denote the set of those positive compact roots which vanish identically on $\nu(\mathfrak{a}_\mathfrak{p})$ and let C' be the complement of C_0 in $P_\mathfrak{k}$. Then if $\rho_\mathfrak{k} = \frac{1}{2}\sum_{\alpha \varepsilon P_\mathfrak{k}} \alpha$ and $\rho_0 = \frac{1}{2}\sum_{\alpha \varepsilon C_0}\alpha$, we have the following result.

LEMMA 25. *Put $\Lambda''_0 = \Lambda_0 + \rho_\mathfrak{k}$. Then if w_0 is the order of the subgroup $\mathfrak{w}_\mathfrak{k}^0$ of $\mathfrak{w}_\mathfrak{k}$ generated by s_α $(\alpha \varepsilon C_0)$,*

$$\chi_{\Lambda_0}(\exp H)\prod_{\alpha \varepsilon C'}(e^{\frac{1}{2}\alpha(H)} - e^{-\frac{1}{2}\alpha(H)})$$
$$= \{w_0 \prod_{\alpha \varepsilon C_0}\rho_0(H_\alpha)\}^{-1}\sum_{s \varepsilon \mathfrak{w}_\mathfrak{k}}\epsilon(s)\{\prod_{\alpha \varepsilon C_0} s\Lambda''_0(H_\alpha)\}e^{s\Lambda''_0(H)}$$

for all $H \varepsilon \nu(\mathfrak{a}_\mathfrak{p})$.

For each $\alpha \varepsilon C'$ we define a holomorphic linear differential operator D_α of order one on the complex Euclidean space $\mathfrak{h}$ such that $D_\alpha\mu = \mu(H_\alpha)$ for every linear function μ on $\mathfrak{h}$. Then the operators D_α obviously commute with each other. Put $D = \prod_{\alpha \varepsilon C'} D_\alpha$. We know (see Weyl [11(a)]) that

$$\chi_{\Lambda_0}(\exp H\prod_{\alpha \varepsilon C'}(e^{\frac{1}{2}\alpha(H)} - e^{-\frac{1}{2}\alpha(H)})\prod_{\alpha \varepsilon C_0}(e^{\frac{1}{2}\alpha(H)} - e^{-\frac{1}{2}\alpha(H)})$$
$$= \sum_{s \varepsilon \mathfrak{w}_\mathfrak{k}}\epsilon(s)e^{s\Lambda''_0(H)} \qquad (H \varepsilon \mathfrak{h}).$$

Hence applying the differential operator D to both sides and evaluating the result at a point H in $\nu(\mathfrak{a}_\mathfrak{p})$, we find that

$$\chi_{\Lambda_0}(\exp H)\prod_{\alpha \varepsilon C'}(e^{\frac{1}{2}\alpha(H)} - e^{-\frac{1}{2}\alpha(H)})(D\Delta_0)_H$$
$$= \sum_{s \varepsilon \mathfrak{w}_\mathfrak{k}}\epsilon(s)\{\prod_{\alpha \varepsilon C_0} s\Lambda''_0(H_\alpha)\}e^{s\Lambda''_0(H)} \qquad H \varepsilon \nu(\mathfrak{a}_\mathfrak{p}))$$

since $\alpha(H)=0$ of $\alpha \varepsilon C_0$ and $H \varepsilon \nu(\mathfrak{a}_\mathfrak{p})$. Here

$$\Delta_0 = \prod_{\alpha \varepsilon C_0} (e^{\alpha/2} - e^{-\alpha/2})$$

and $(D\Delta_0)_H$ denotes the value of $D\Delta_0$ at H. But it follows from well-known arguments (Weyl [11(a)]) that[9]

$$\Delta_0 = \sum_{s \varepsilon \mathfrak{w}_\mathfrak{k}^0} \epsilon_0(s) e^{s\rho_0}$$

where $\epsilon_0(s) = \pm 1$ and is determined by the rule

$$\prod_{\alpha \varepsilon C_0} (e^{\frac{1}{2}s\alpha} - e^{-\frac{1}{2}s\alpha}) = \epsilon_0(s)\Delta_0 \qquad (s \varepsilon \mathfrak{w}_\mathfrak{k}^0).$$

Hence

$$D\Delta_0 = \sum_{s \varepsilon \mathfrak{w}_\mathfrak{k}^0} \epsilon_0(s) \{\prod_{\alpha \varepsilon C_0} s\rho_0(H_\alpha)\} e^{s\rho_0}.$$

But it is obvious that $\epsilon_0(s) = (-1)^q$ $(s \varepsilon \mathfrak{w}_\mathfrak{k}^0)$ if q is the number of negative roots among $s\alpha$ $(\alpha \varepsilon C_0)$. Therefore

$$\prod_{\alpha \varepsilon C_0} s\rho_0(H_\alpha) = \epsilon_0(s) \prod_{\alpha \varepsilon C_0} \rho_0(H_\alpha)$$

and

$$(D\Delta_0)_H = w_0 \prod_{\alpha \varepsilon C_0} \rho_0(H_\alpha) \qquad (H \varepsilon \nu(\mathfrak{a}_\mathfrak{p}))$$

since $\rho_0(H)=0$. Moreover if $\beta \varepsilon C_0$, not all the roots $s_\beta\alpha (\alpha \varepsilon C_0)$ are positive since $s_\beta\beta = -\beta$. This shows that $\sum_{\alpha \varepsilon C_0} s_\beta\alpha < \sum_{\alpha \varepsilon C_0} \alpha$ and therefore $s_\beta\rho_0 < \rho_0$. Since this implies that $\rho_0(H_\beta) > 0$, the lemma now follows.

Now again for any point $H = \sum_{i=1}^{s} t_i(X_{\gamma_i} + X_{-\gamma_i})$ $(t_i \varepsilon R)$ in $\mathfrak{b}$ put $H' = \sum_i \log(\cosh t_i) H_{\gamma_i}$. Then if $\rho = \frac{1}{2}\sum_{\alpha \varepsilon F} \alpha$ and $\Lambda' = \Lambda + \rho$, $\Lambda' = \Lambda''_0 + \lambda + \rho_+$. Moreover since $\lambda(H_\alpha) = 0$ for every compact root α [5(e), Lemma 13], it follows that $s\lambda = \lambda$ for all $s \varepsilon \mathfrak{w}_\mathfrak{k}$. Similarly it follows from Lemma 10 of [5(e)] that $s\rho_+ = \rho_+$ $(s \varepsilon \mathfrak{w}_\mathfrak{k})$ and therefore $\rho_+(H_\alpha) = 0$ if $\alpha \varepsilon P_\mathfrak{k}$. Hence we conclude from Lemmas 23 and 25 that

$$\chi_{\Lambda_0}(\exp 2H') \prod_{i=1}^{s} (\cosh t_i)^{2\lambda(H_{\gamma_i})} |D(\exp H)|^{\frac{1}{2}}$$
$$= 2^p\{w_0 \prod_{\alpha \varepsilon C_0} \rho_0(H_\alpha)\}^{-1} \prod_i (\sinh t_i)$$
$$\times \sum_{\tau \varepsilon \mathfrak{w}_\mathfrak{k}} \epsilon(\tau) \{\prod_{\alpha \varepsilon C_0} \tau\Lambda'(H_\alpha)\} \prod_i (\cosh t_i)^{2\tau\Lambda'(H_{\gamma_i})-1}.$$

[9] We have to apply here Weyl's result to the subalgebra $\mathfrak{a}_\mathfrak{k} + \sum_{\alpha \varepsilon C_0} (CX_\alpha + CX_{-\alpha})$ of $\mathfrak{g}$ which is obviously reductive [7].

Put $y_i = (\cosh t_i)^{-1}$. Then $\mathfrak{b}$ corresponds to the region defined by $0 < y_i < 1$ and $y_i > y_j$ if $\gamma_i \prec \gamma_j$ $(1 \leqq i < j \leqq s)$. Let us denote this region by $\mathfrak{b}_y$. Then if $dy = dy_1 dy_2 \cdots dy_s$, it is clear that

$$\int_{\mathfrak{b}} \chi_{\Lambda_0}'(\exp 2H') e^{2\lambda(\Gamma(\exp H))} \,|\, D(\exp H)\,|^{\frac{1}{2}}\, dt$$
$$= 2^p\{w_0 \prod_{\alpha \in C_0} \rho_0(H_\alpha)\}^{-1} \int_{\mathfrak{b}_y} \{ \sum_{\tau \in \mathfrak{w}_{\mathfrak{l}}} \epsilon(\tau) (\prod_{\alpha \in C_0} \tau\Lambda'(H_\alpha)) \prod_i y_i^{-2\tau\Lambda'(H\gamma_i)-1} \} dy.$$

Put $y_{ij} = (\cosh t_{ij})^{-1}$ $(1 \leqq i \leqq r, 1 \leqq j \leqq s_i)$ in the notation of Section 8. Then the region $\mathfrak{b}_y$ is defined by the inequalities

$$1 > y_{i1} > y_{i2} > \cdots > y_{is_i} > 0 \qquad (1 \leqq i \leqq r)$$

and an elementary computation shows that if q_{ij} are *positive* real numbers

$$\int_{\mathfrak{b}_y} \prod_{i,j} y_{ij}^{q_{ij}-1}\, dy = \prod_{1 \leqq i \leqq r} \prod_{1 \leqq k \leqq s_i} (\sum_{k \leqq j \leqq s_i} q_{ij})^{-1}.$$

We shall use this formula to determine the value of our integral.

Let $\mathfrak{F}$ denote the space of linear functions on $\mathfrak{h}$. By a rational function ω on $\mathfrak{F}$, we mean an element of the quotient field of the ring of polynomial functions on $\mathfrak{F}$ (see [5(b), p. 194]). We say that ω is defined at a point $\mu \in \mathfrak{F}$, if it can be written in the form $\omega = f/g$ where f and g are two polynomial functions on $\mathfrak{F}$ and $g(\mu) \neq 0$. It is obvious that in that case the ratio $f(\mu)/g(\mu)$ depends only on ω (and not on the choice of f and g). This ratio is called the value of ω at μ and denoted by $\omega(\mu)$. We shall need the following simple lemma on rational functions.

LEMMA 26. *Let f and $g \neq 0$ be two polynomial functions on F and let $\omega = f/g$. Suppose there exists a base $\lambda_1, \cdots, \lambda_l$ for $\mathfrak{F}$ over C and an integer N_0 with the following property. For any given integers $m_1, \cdots, m_l \geqq N_0$, $\omega(m_1\lambda_1 + \cdots + m_l\lambda_l) = 0$ provided $g(m_1\lambda_1 + \cdots + m_l\lambda_l) \neq 0$. Then $f = 0$.*

For otherwise $fg \neq 0$ and therefore we can choose integers $m_1, \cdots, m_l \geqq N_0$ such that $f(m_1\lambda_1 + \cdots + m_l\lambda_l) g(m_1\lambda_1 + \cdots + m_l\lambda_l) \neq 0$ (see Lemma 32 of [5(a)]). Then obviously $\omega(m_1\lambda_1 + \cdots + m_l\lambda_l) \neq 0$ and we get a contradiction.

COROLLARY. *Suppose ω_1, ω_2 are two rational functions on $\mathfrak{F}$ such that*

$$\omega_1(m_1\lambda_1 + \cdots + m_l\lambda_l) = \omega_2(m_1\lambda_1 + \cdots + m_l\lambda_l)$$

for all integral values of $m_1, \cdots, m_l \geqq N_0$ whenever both sides are defined. Then $\omega_1 = \omega_2$.

Let $\omega_1 = f_1/g_1$, $\omega_2 = f_2/g_2$ where f_1, f_2, g_1, g_2 are polynomial functions and $g_1 \neq 0$, $g_2 \neq 0$. Then $\omega_1 - \omega_2 = (g_2 f_1 - g_1 f_2)/g_1 g_2$. Since ω_1 and ω_2 are both defined at any point where $g_1 g_2$ does not vanish, it follows from the above lemma that $g_2 f_1 - g_1 f_2 = 0$ and therefore $\omega_1 = \omega_2$.

For any polynomial function f and $\tau \in \mathfrak{w}_{\mathfrak{k}}$, we denote by f^τ the polynomial function $\mu \rightarrow f(\tau^{-1}\mu)$ $(\mu \in \mathfrak{F})$. Now define a polynomial function F_0 as follows:

$$F_0(\mu) = \prod_{1 \leqq i \leqq s} \prod_{1 \leqq k \leqq s_i} \Big(\sum_{k \leqq j \leqq s_i} q_{ij}(\mu) \Big) \qquad (\mu \in \mathfrak{F})$$

where $q_{ij}(\mu) = \mu(H_{\gamma_p})$ if $\beta_{ij} = \gamma_p$ $(1 \leqq p \leqq s)$ in the notation of Section 8. Similarly let f_0, g_τ and F_τ $(\tau \in \mathfrak{w}_{\mathfrak{k}})$ be the polynomial functions given by

$$f_0(\mu) = \{w_0 \prod_{\alpha \in C_0} \rho_0(H_\alpha)\} \prod_{\alpha \in P_{\mathfrak{k}}} \{(\mu(H_\alpha) + \rho(H_\alpha))/\rho(H_\alpha)\}^2$$

$$g(\mu) = \prod_{\alpha \in C_0} \{\mu(\tau H_\alpha) + \rho(\tau H_\alpha)\}$$

$$F_\tau(\mu) = F_0{}^\tau(-2(\mu + \rho)) \qquad (\mu \in \mathfrak{F})$$

and put

$$\omega = 2^p \sum_{\tau \in \mathfrak{w}_{\mathfrak{k}}} \epsilon(\tau) g_\tau / f_0 F_\tau$$

where $p = 2 \sum_{1 \leqq i \leqq s} \rho_+(H_{\gamma_i})$. Then it follows from what we have proved above that

$$\int_{G_0} |\psi_\Lambda(x)|^2 \, d\bar{x} = 2^p w c N^{-1} \{f_0(\Lambda)\}^{-1} \int_{\mathfrak{h}_y} \{ \sum_{\tau \in \mathfrak{w}_{\mathfrak{k}}} \epsilon(\tau) g_\tau(\Lambda) \prod_i y_i^{-2\tau\Lambda'(H_{\gamma_i})-1} \} \, dy.$$

Now let us suppose that $\Lambda'(H_\gamma) < 0$ for every totally positive root γ. Then since any $\tau \in \mathfrak{w}_{\mathfrak{k}}$ permutes totally positive roots among themselves [5(e), Lemma 10], it follows that $\tau\Lambda'(H_\gamma) < 0$ $(\gamma \in P_+)$. Hence the above-mentioned formula is applicable to each term of the integral on the right and we see that the rational function ω is defined at Λ and

$$\int_{G_0} |\psi_\Lambda(x)|^2 \, d\bar{x} = w c N^{-1} \omega(\Lambda).$$

Thus we have the following result.

LEMMA 27. *Let $\mathfrak{F}_G(P)$ denote the set of all linear functions on* Λ *satisfying the following two conditions:*

(1) *Λ is real and $\Lambda(H_\alpha)$ is a non-negative integer for every positive compact root α.*

(2) *$\Lambda(H_\gamma) + \rho(H_\gamma) < 0$ for every totally positive root γ. Then there*

exists a uniquely determined rational function ω *on* $\mathfrak{F}$ *which is defined everywhere on* $\mathfrak{F}_G(P)$ *and such that*

$$\int_{G_0} |\psi_\Lambda(x)|^2 \, d\bar{x} = wcN^{-1}\omega(\Lambda)$$

for all $\Lambda \in \mathfrak{F}_G(P)$.

We have only to show that ω is unique. Let $(\alpha_1, \cdots, \alpha_l)$ be a fundamental system of positive roots and suppose $(\alpha_1, \cdots, \alpha_m)$ are all the totally positive roots in this system. Define l linear functions Λ_i $(1 \leqq i \leqq l)$ on $\mathfrak{h}$ by the conditions $\Lambda_i(H_{\alpha_j}) = \delta_{ij}$ $1 \leqq i, j \leqq l$. Then if $\lambda_+ = \Lambda_1 + \cdots + \Lambda_m$, it follows from Lemma 13 of [5(e)] that $\lambda_+(H_\gamma) > 0$ for $\gamma \in P_+$ and $\lambda_+(H_\alpha) = 0$ for $\alpha \in P_{\mathfrak{k}}$. Hence it is obvious that we can find an integer $n \geqq 2$ such that $\Lambda_i - n\lambda_+ \in \mathfrak{F}_G(P)$ $(1 \leqq i \leqq l)$. Therefore if $\lambda_i = \Lambda_i - n\lambda_+$, $(\lambda_1, \cdots, \lambda_l)$ is a base for $\mathfrak{F}$ over C and $m_1\lambda_1 + \cdots + m_l\lambda_l \in \mathfrak{F}_G(P)$ for every set of positive integers $(m_1, \cdots, m_l)$. The uniqueness of ω now follows immediately from the Corollary to Lemma 26.

We shall now determine ω in another way. Let $\mathfrak{F}_U$ denote the set of all real linear functions Λ on $\mathfrak{h}$ such that $\Lambda(H_\beta)$ is a non-negative integer for every positive root β. For a fixed $\Lambda \in \mathfrak{F}_U$, let σ denote the irreducible complex representation (see [5(f), § 6]) of G_c on a finite-dimensional Hilbert space V with the highest weight Λ. We assume that σ is unitary on U. ϕ_0, being a unit vector in V belonging to the weight Λ, we consider the function

$$\psi_\Lambda^*(x) = (\phi_0, \sigma(x)\phi_0) \qquad (x \in G_c).$$

Since V is irreducible under $\sigma(U)$, it follows from the Schur orthogonality relations for the compact group U, that

$$\int_U |\psi_\Lambda^*(u)|^2 \, du = (\dim V)^{-1}.$$

On the other hand if we normalize the various Haar measures according to Lemma 22 and put $\Lambda' = \Lambda + \rho$, we get

$$\int_U |\psi_\Lambda^*(u)|^2 \, du = w^*N^{-1}\Big(\int_{\mathfrak{A}^*} |D(a^*)|^{\frac{1}{2}} \, da^*\Big)^{-1}$$
$$\times \int_{B^*} |D(a^*)|^{\frac{1}{2}} \, da^* \int_{K_0 \times K_0} |\psi_\Lambda^*(ka^*k')|^2 \, dk dk'$$

where B^* is defined as in Lemma 24 and we make use of the fact (see the Corollary to Lemma 24) that B^* and B both have N connected components. Now let $H = \sum_i t_i(X_{\gamma_i} + X_{-\gamma_i}) \in \mathfrak{b}^*$ $(t_i \in R)$ and $a^* = e(H)$. Then if we put

$$H^* = \sum_i \log(\cos t_i) H_{\gamma_i}$$

and $h(a^*) = \exp H^* \varepsilon A_c$, we know from Lemma 20 that

$$a^* = \zeta h(a^*) \xi$$

where $\zeta \varepsilon \mathfrak{P}_c^-$ and $\xi \varepsilon \mathfrak{P}_c^+$. Therefore we conclude in the same way as before, that

$$(\phi_0, \sigma(ka^*k')\phi_0) = (\phi_0, \sigma(kh(a^*)k')\phi_0)$$

and therefore

$$\int_{K_0 \times K_0} |(\phi_0, \sigma(ka^*k')\phi_0)|^2 \, dkdk' = \int_{K_0 \times K_0} |(\phi_0, \sigma(kh(a^*)k')\phi_0)|^2 \, dkdk'.$$

On the other hand if V_0 is the subspace of V spanned by $\sigma(k)\phi_0$ $(k \varepsilon K_c)$, we prove exactly as before that the corresponding representation σ_0 of K_c on V_0 is irreducible and has the highest weight Λ. Then if χ_Λ denotes the character of σ_0 we have

$$\int_{K_0 \times K_0} |(\phi_0, \sigma(ka^*k')\phi_0)|^2 \, dkdk' = (\dim V_0)^{-2} \chi_\Lambda((h(a^*))^2).$$

Therefore

$$\int_{B^*} |D(a^*)|^{\frac{1}{2}} da^* \int_{K_0 \times K_0} |\psi_\Lambda^*(ka^*k')|^2 \, dkdk'$$

$$= (\dim V_0)^{-2} \int_{B^*} \chi_\Lambda((h(a^*))^2) |D(a^*)|^{\frac{1}{2}} da^*$$

$$= (\dim V_0)^{-2} c \int_{\mathfrak{b}^*} \chi_\Lambda(\exp 2H^*) |D(e(H))|^{\frac{1}{2}} dt$$

where c and dt have the same meaning as before. But from Lemmas 23 and 25 it follows that[10]

$$\chi_\Lambda(\exp 2H^*) |D(e(H))|^{\frac{1}{2}}$$
$$= (-1)^p 2^p \{w_0 \prod_{\alpha \varepsilon C_0} \rho(H_\alpha)\}^{-1} \{\prod_{1 \leqq i \leqq s} \sin t_i\}$$
$$\times \sum_{\tau \varepsilon \mathfrak{w}_{\mathfrak{k}}} \{\epsilon(\tau) g_\tau(\Lambda) \prod_{1 \leqq i \leqq s} (\cos t_i)^{2\tau\Lambda'(H\gamma_i) - 1}\}$$

where $\Lambda' = \Lambda + \rho$. Now put $y_i = \cos t_i$. Then $\mathfrak{b}^*$ corresponds to the region $\mathfrak{b}_y$ defined as before by $0 < y_i < 1$ and $y_i > y_j$ if $\gamma_i \prec \gamma_j$ $(1 \leqq i < j \leqq s)$.

---

[10] The proof of Lemma 25 made no use of the fact that $\Lambda_0(H_\alpha) = 0$ for every totally positive root α lying in the fundamental system $(\alpha_1, \cdots, \alpha_l)$ of positive roots. Hence this lemma is applicable also to χ_Λ.

Therefore if we use Weyl's formula [11(a)] for $\dim V_0$ and recall that $\rho(H_\alpha)=\rho_{\mathfrak{k}}(H_\alpha)$ $(\alpha\in P_{\mathfrak{k}})$, we get

$$(\dim V_0)^{-2}\int_{\mathfrak{h}^*}\chi_\Lambda(\exp 2H^*)\,|D(e(H))|^{\frac{1}{2}}\,dt$$
$$=(-1)^p\{2^p/f_0(\Lambda)\}\int_{\mathfrak{h}_v}\{\sum_{\tau\varepsilon\mathfrak{w}_{\mathfrak{k}}}\epsilon(\tau)g_\tau(\Lambda)\prod_{1\leq i\leq s}y_i^{2\tau\Lambda'(H_{\gamma_i})-1}\}\,dy.$$

We have seen that $\tau\varepsilon\mathfrak{w}_{\mathfrak{k}}$ permutes the totally positive roots among themselves and if $\gamma\varepsilon P_+$, $\Lambda'(H_\gamma)>0$ since $\Lambda(H_\gamma)\geqq 0$ and $\rho(H_\gamma)>0$ (see Weyl [11(b)]). Therefore $2\tau\Lambda'(H_{\gamma_i})>0$ and it follows that the right hand side of the above equation is

$$=(-1)^p2^p\sum_{\tau\varepsilon\mathfrak{w}_{\mathfrak{k}}}\epsilon(\tau)g_\tau(\Lambda)\{f_0(\Lambda)F_0{}^\tau(2\Lambda')\}^{-1}.$$

Put $\Lambda^*=-(\Lambda+2\rho)$. Then it is obvious from the definition of F_0, f_0, F_τ and g_τ $(\tau\varepsilon\mathfrak{w}_{\mathfrak{k}})$ that

$$F_\tau(\Lambda^*)=F_0{}^\tau(2\Lambda'),f_0(\Lambda^*)=f_0(\Lambda),g_\tau(\Lambda^*)=(-1)^{r_0}g_\tau(\Lambda)$$

where r_0 is the number of roots in C_0. Hence

$$\int_U|\psi_{\Lambda^*}(u)|^2\,du=w^*cN^{-1}(\int_{\mathfrak{A}^*}|D(a^*)|^{\frac{1}{2}}\,da^*)^{-1}(-1)^{p+r_0}\omega(\Lambda^*).$$

But on the other hand by Weyl's formula [11(a)],

$$\int_U|\psi_{\Lambda^*}(u)|^2\,du=(\dim V)^{-1}=\{\prod_{\beta\varepsilon P}\Lambda'(H_\beta)/\rho(H_\beta)\}^{-1}$$

since Λ is the highest weight of σ. This proves that

$$\omega(\Lambda^*)=(-1)^{p+r_0+m}(N/cw^*)\int_{\mathfrak{A}^*}|D(a^*)|^{\frac{1}{2}}\,da^*$$
$$\times\{\prod_{\beta\varepsilon P}(\Lambda^*(H_\beta)+\rho(H_\beta))/\rho(H_\beta)\}^{-1}$$

where m is the total number of positive roots. Let $\mathfrak{F}_U{}^*$ denote the set of all linear functions of the form $\Lambda^*=-(\Lambda+2\rho)$ $(\Lambda\varepsilon\mathfrak{F}_U)$. Define the linear functions Λ_i $1\leqq i\leqq l$ as in the proof of Lemma 27. Then if $\lambda_1=-(\Lambda_1+2\rho)$ and $\lambda_i=-\Lambda_i$ $2\leqq i\leqq l$, $(\lambda_1,\cdots,\lambda_l)$ is a base for $\mathfrak{F}$ and $m_1\lambda_1+\cdots+m_l\lambda_l\varepsilon\mathfrak{F}_U{}^*$ for positive integers $(m_1,\cdots,m_l)$. Therefore in view of the Corollary to Lemma 26, we can conclude that

$$\omega(\mu)=(-1)^{p+r_0+m}(N/cw^*)\int_{\mathfrak{A}^*}|D(a^*)|^{\frac{1}{2}}\,da^*$$
$$\times\{\prod_{\beta\varepsilon P}(\mu(H_\beta)+\rho(H_\beta))/\rho(H_\beta)\}^{-1}$$

$(\mu \varepsilon \mathfrak{F})$, both sides being defined and equal whenever at least one of them is defined. Hence in particular if $\Lambda \varepsilon \mathfrak{F}_G(P)$, it follows from Lemma 27 that

$$\int_{G_0} |\psi_\Lambda(x)|^2 d\bar{x} = (-1)^{p+r_0+m}(w/w^*) \int_{\mathfrak{A}^*} |D(a^*)|^{\frac{1}{2}} da^* \times \{ \prod_{\beta \varepsilon P} \Lambda'(H_\beta)/\rho(H_\beta) \}^{-1}$$

where $\Lambda' = \Lambda + \rho$. On the other hand if q is the number of totally positive roots, $\prod_{\beta \varepsilon P} \Lambda'(H_\beta)$ has obviously the same sign as $(-1)^q$. Therefore since the left side is positive and $\rho(H_\beta) > 0$ for every positive root β, $(-1)^{p+r_0+m} = (-1)^q$. Thus we have obtained the following result.

LEMMA 28. *If* $\Lambda \varepsilon \mathfrak{F}_G(P)$,

$$\int_{G_0} |\psi_\Lambda(x)|^2 d\bar{x} = (-1)^q(w/w^*) \int_{\mathfrak{A}^*} |D(a^*)|^{\frac{1}{2}} da^* \times \{ \prod_{\beta \varepsilon P} (\Lambda(H_\beta) + \rho(H_\beta))/\rho(H_\beta) \}^{-1}.$$

Although our definition of $\mathfrak{a}_{\mathfrak{p}_0}$ depends on the order (which we have so far assumed to be fixed) on the space $\mathfrak{F}_R$ of real linear functions on $\mathfrak{h}$, we know from Lemma 22 that the above normalization of the measure $d\bar{x}$ and the numbers w, w^*, $\int_{\mathfrak{A}^*} |D(a^*)|^{\frac{1}{2}} da^*$ are actually independent of this order. Therefore if we normalize $d\bar{x}$ in such a way that

$$\int_{G_0} |\psi_\Lambda(x)|^2 d\bar{x} = |\prod_{\beta \varepsilon P} \Lambda(H_\beta) + \rho(H_\beta)/\rho(H_\beta)|^{-1}$$

for all $\Lambda \varepsilon \mathfrak{F}_G(P)$, this new normalization is also independent of the particular order in $\mathfrak{F}_R$. However since the definition of the function ψ_Λ depends on this order, we shall now denote it by $\psi_\Lambda{}^P$. Then our result may be stated as follows.

THEOREM 4. *It is possible to normalize the Haar measure on $d\bar{x}$ on G_0 in such a way that the following condition is fulfilled. Let P be the set of all positive roots under any given order $\mathfrak{F}_R$ and suppose every noncompact root in P is totally positive. Let $\rho = \frac{1}{2}\sum_{\beta \varepsilon P} \beta$ and let $\mathfrak{F}_G(P)$ denote the set of all real linear functions Λ on $\mathfrak{h}$ which satisfy the following two conditions:*

(1) $\Lambda(H_\alpha)$ *is a non-negative integer for every positive compact root* α.

(2) $\Lambda(H_\beta) + \rho(H_\beta) < 0$ *for noncompact positive root* β.

Then

$$\int_{G_0} | \psi_\Lambda{}^P(x) |^2 d\bar{x} = | \prod_{\beta \varepsilon P} \Lambda(H_\beta) \rho(H_\beta)/\rho(H_\beta) |^{-1}$$

for all $\Lambda \varepsilon \mathfrak{F}_G(P)$.

Now we return again to our fixed order in $\mathfrak{F}_R$. Let Λ be a real linear function on $\mathfrak{h}$ such that $\Lambda(H_\alpha)$ is a non-negative integer for every compact positive root α. Define $\psi_\Lambda(x)$ $(x \varepsilon G)$ as in the beginning of this section.

LEMMA 29. $\int_{G_0} | \psi_\Lambda(x) |^2 d\bar{x} < \infty$ *if and only if* $\Lambda \varepsilon \mathfrak{F}_G(P)$.

Define λ_+ as in the proof of Lemma 27. Then, as we have seen, $\lambda_+(H_\gamma) > 0$ for $\gamma \varepsilon P_+$ and $\lambda_+(H_\alpha) = 0$ if $\alpha \varepsilon P_{\mathfrak{k}}$. Put

$$t_0 = \max_{\gamma \varepsilon P_+} \{\Lambda(H_\gamma) + \rho(H_\gamma)\}/\lambda_+(H_\gamma).$$

Then it is obvious that $\Lambda - t\lambda_+ \varepsilon \mathfrak{F}_G(P)$ $(t \varepsilon R)$ if and only if $t > t_0$ and therefore $t_0 \geqq 0$ if and only if $\Lambda \notin \mathfrak{F}_G(P)$. Moreover if $\Lambda_t = \Lambda - t\lambda_+$, it is clear that

$$\psi_{\Lambda_t}(x) = \psi_\Lambda(x) e^{-t\lambda_-(\Gamma(x))} \qquad (x \varepsilon G).$$

But if $H = \sum_{1 \leqq i \leqq s} t_i(X_{\gamma_i} + X_{-\gamma_i})$ $(t_i \varepsilon R)$ we have seen above that

$$\lambda_+(\Gamma(\exp H)) = \sum_{1 \leqq i \leqq s} (\log \cosh t_i) \lambda_+(H_{\gamma_i}).$$

Since $\cosh t_i \geqq 1$ and $\lambda_+(H_{\gamma_i}) \geqq 0$, it follows that $\lambda_+(\Gamma(a)) \geqq 0$ $(a \varepsilon \mathfrak{A})$. Therefore it is obvious from Lemma 22, that

$$\int_{G_0} | \psi_\Lambda(x) |^2 d\bar{x} = \int_{G_0} | \psi_{\Lambda_t}(x) e^{t\lambda_+(\Gamma(x))} |^2 d\bar{x} \geqq \int_{G_0} | \psi_{\Lambda_t}(x) |^2 d\bar{x}$$

provided $t \geqq 0$. Now suppose $\Lambda \notin \mathfrak{F}_G(P)$ so that $t_0 \geqq 0$. Put $t = t_0 + \epsilon$ where ϵ is positive. Then, in view of Theorem 4, we get

$$\int_{G_0} \psi_\Lambda(x) |^2 d\bar{x} \geqq | \prod_{\beta \varepsilon P} \{\Lambda(H_\beta) + \rho(H_\beta) - (t_0 + \epsilon)\lambda_+(H_\beta)\}/\rho(H_\beta) |^{-1}.$$

But it follows from the definition of t_0 that as $\epsilon \to 0$ the right side tends to infinity. Therefore

$$\int_{G_0} | \psi_\Lambda(x) |^2 d\bar{x} = \infty.$$

Conversely if $\Lambda \varepsilon \mathfrak{F}_G(P)$ we know from Theorem 4 that

$$\int_{G_0} | \psi_\Lambda(x) |^2 d\bar{x} < \infty.$$

Thus the lemma is proved.

We shall now consider the question of the integrability of the function ψ_Λ. Let us recall that $\rho_+=\frac{1}{2}\sum_{\beta\varepsilon P_+}\beta$.

LEMMA 30. *Let Λ be a linear function on $\mathfrak{h}$ satisfying the following two conditions:*

(1) *$\Lambda(H_\alpha)$ is a non-negative integer for every compact positive root α.*

(2) *$\Lambda(H_\beta)+\rho(H_\beta)<1-2\rho_+(H_\beta)$ for every noncompact positive root β.*

Then

$$\int_{G_0}|\psi_\Lambda(x)|^2\,d\bar{x}<\infty.$$

We use the notation which was introduced at the beginning of this section. In view of Lemma 22, it is enough to prove that

$$\int_B|D(a)|^{\frac{1}{2}}da\int_{K_0\times K_0}|(\phi_0,\sigma(kak')\phi_0)|\,e^{\lambda(\Gamma(a))}\,dkdk'<\infty.$$

But it follows from the Schwartz inequality that

$$\int_{K_0\times K_0}|(\phi_0,\sigma(kak')\phi_0)|\,dkdk'\leqq\{\int_{K_0\times K_0}|(\phi_0,\sigma(kak')\phi_0)|^2\,dkdk'\}^{\frac{1}{2}}$$
$$=(\dim V_0)^{-1}\{\mathrm{Sp}\,\sigma_0((h(a))^2)\}^{\frac{1}{2}}$$

as we have seen before. Since $\sigma_0(h(a))$ is obviously a positive definite self-adjoint transformation on V_0, it is clear that

$$\mathrm{Sp}\,\sigma_0((h(a))^2)\leqq\{\mathrm{Sp}\,\sigma_0(h(a))\}^2.$$

Therefore

$$\int_{K_0\times K_0}|(\phi_0,\sigma(kak')\phi_0)|\,dkdk'\leqq(\dim V_0)^{-1}\chi_{\Lambda_0}(h(a)).$$

So it would be sufficient to prove that

$$\int_B\chi_{\Lambda_0}(h(a))e^{\lambda(\Gamma(a))}\,|D(a)|^{\frac{1}{2}}da<\infty.$$

For any point

$$H=\sum_{1\leqq i\leqq s}t_i(X_{\gamma_i}+X_{-\gamma_i})\ \ (t\,\varepsilon\,R)\text{ in }\mathfrak{b}\text{ we put }H'=\sum_{1\leqq i\leqq s}(\log\cosh t_i)H_{\gamma_i}$$

as before. Then if $a=\exp H$, $h(a)=\exp H'$ and therefore it follows from Lemmas 23 and 25 that

$$\chi_{\Lambda_0}(h(a))e^{\lambda(\Gamma(a))}\,|D(a)|^{\frac{1}{2}}$$
$$=c_1\frac{\Delta(2H')}{\Delta(H')}\{\sum_{\tau\,\varepsilon\,\mathfrak{w}_\mathfrak{f}}\epsilon(\tau)g_\tau(\Lambda)\prod_{i=1}^{s}(\cosh t_i)^{\tau\Lambda'(H\gamma_i)+\rho_+(H\gamma_i)-1}\}\prod_{i=1}^{s}(\sinh t_i)$$

where $\Lambda' = \Lambda + \rho$, c_1 is a positive constant and g_τ is defined as before. Now

$$\frac{\Delta(2H')}{\Delta(H')} = \prod_{\alpha \in C'} (e^{\frac{1}{2}\alpha(H')} + e^{-\frac{1}{2}\alpha(H')})$$

and therefore if r' is the number of roots in C', it follows from Lemma 16 that

$$0 < \frac{\Delta(2H')}{\Delta(H')} \leqq 2^{r'} \prod_{\alpha \in C'} e^{\frac{1}{2}|\alpha(H')|}$$

$$= 2^{r'} \prod_{i=1}^{s} (\cosh t_i)^{\frac{1}{2}r_i} \prod_{1 \leqq i < j \leqq s} (\cosh t_j / \cosh t_i)^{\frac{1}{2}r_{ij}}$$

in the notation of Section 6. (Here we have to make use of the fact that $t_j > t_i$ on $\mathfrak{b}$ if $r_{ij} > 0$ $(1 \leqq i < j \leqq s)$). But since

$$r_i - \sum_{i < j \leqq s} r_{ij} + \sum_{1 \leqq j < i} r_{ji} \leqq 2\rho_+(H_{\gamma_i}) - 2$$

from Lemma 18, we conclude that

$$0 < \{\Delta(2H')/\Delta(H')\} \leqq 2^{r'} \prod_{i=1}^{s} (\cosh t_i)^{\rho_+(H\gamma_i)-1}$$

and therefore if $c_2 = 2^{r'} c_1$,

$$\chi_{\Lambda_0}(h(a)) e^{\lambda(\Gamma(a))} \, |D(a)|^{\frac{1}{2}}$$

$$\leqq c_2 \{ \sum_{\tau \in \mathfrak{w}_{\mathfrak{k}}} \epsilon(\tau) g_\tau(\Lambda) \prod_i (\cosh t_i)^{\tau\Lambda'(H\gamma_i) + 2\rho_+(H\gamma_i) - 2} \} \prod_i \sinh t_i.$$

Now put $y_i = (\cosh t_i)^{-1}$, $dy = dy_1 \cdots dy_s$ and define $\mathfrak{b}_y$ as before. Then

$$\int_{\mathfrak{b}} \prod_i (\cosh t_i)^{\tau\Lambda'(H\gamma_i) + 2\rho_+(H\gamma_i) - 2} \prod_i (\sinh t_i) \, dt$$

$$= \int_{\mathfrak{b}_y} \prod_i y_i^{-\tau\Lambda'(H\gamma_i) - 2\rho_+(H\gamma_i)} \, dy \qquad (\tau \in \mathfrak{w}_{\mathfrak{k}}).$$

But we have seen earlier that $\tau\rho_+ = \rho_+$ and therefore

$$\tau\Lambda'(H_{\gamma_i}) + 2\rho_+(H_{\gamma_i}) = \Lambda'(H_\gamma) + 2\rho_+(H_\gamma)$$

where $\gamma = \tau^{-1}\gamma_i$. Hence it follows from our hypothesis on Λ that

$$-\tau\Lambda'(H_{\gamma_i}) - 2\rho_+(H_{\gamma_i}) + 1 > 0$$

and therefore

$$\int_{\mathfrak{b}_y} \prod_i y_i^{-\tau\Lambda'(H\gamma_i) - 2\rho_+(H\gamma_i)} \, dy < \infty.$$

This shows that

$$\int_B \chi_{\Lambda_0}(h(a)) e^{\lambda(\Gamma(a))} \, |D(a)|^{\frac{1}{2}} \, da < \infty$$

and so the lemma is proved.

10. Integrable and square-integrable representations. Suppose Λ is a real linear function on $\mathfrak{h}$ which satisfies the first condition of Lemma 30. We define the Hilbert space $\mathfrak{H}_\Lambda$ as in Theorem 2 of [5(f)] corresponding to the function $\mu = 1$ on G_0. We know from Lemma 14 of [5(f)] and Lemma 29 that $\mathfrak{H}_\Lambda \neq 0$ if and only if $\Lambda \varepsilon \mathfrak{F}_G(P)$. So now let us assume that $\Lambda \varepsilon \mathfrak{F}_G(P)$ and let π_Λ denote the representation of G on $\mathfrak{H}_\Lambda$ (see [5(f), Theorem 2]). Then π_Λ is irreducible and unitary and

$$(\psi_\Lambda, \pi_\Lambda(x)) = \psi_\Lambda(x) \| \psi_\Lambda \|^2 \qquad (x \varepsilon G)$$

from the Corollary to Theorem 2 of [5(f)]. (Here we use the usual notation for the norm and the scalar product in $\mathfrak{H}_\Lambda$). Since $\psi_\Lambda \neq 0$ and

$$\int_{G_0} | \psi_\Lambda(x) |^2 d\bar{x} = \| \psi_\Lambda |^2 < \infty,$$

this prove that π_Λ is square-integrable (see Section 3). Moreover if Λ satisfies the second condition of Lemma 30, we prove in the same way that π_Λ is integrable. Combining these results with Lemma 19 of [5(e)], we get Theorem 3 of [5(e)] which was stated there without proof.

It is clear from the definition of $\mathfrak{H}_\Lambda$ that every element $\phi \varepsilon \mathfrak{H}_\Lambda$ is compeletly determined by its restriction on G and in fact

$$\| \phi \|^2 = \int_{G_0} | \phi(x) |^2 d\bar{x}.$$

Therefore since $\psi_\Lambda \varepsilon \mathfrak{H}_\Lambda$ and π_Λ is irreducible, $\mathfrak{H}_\Lambda$ may be identified with the completion (under the above norm) of the space spanned by the right translates of the function ψ_Λ under G. If we use the normalization of Theorem 4 for the Haar measure on G_0 and denote by d_Λ the formal degree of π_Λ (corresponding to the kernel Z_0 of the mapping $x \to \bar{x}$ of G on G_0 (see Section 3)), it follows from Theorem 1 that

$$d_\Lambda^{-1} \| \psi_\Lambda \|^2 = \int_{G_0} | (\psi_\Lambda, \pi_\Lambda(x)\psi_\Lambda) |^2 d\bar{x} = \| \psi_\Lambda \|^4$$

and therefore

$$d_\Lambda = \| \psi_\Lambda \|^{-2} = \prod_{\beta \varepsilon P} | (\Lambda(H_\beta) + \rho(H_\beta))/\rho(H_\beta) |$$

from Theorem 4. This is the formula for the formal degree in terms of the highest weight Λ of the representation. It is substantially the same as the corresponding formula of Weyl [11(a)] for a compact semisimple group.

11. Similarity with finite-dimensional representations. In order to simplify matters let us suppose in this section that G is simple but not compact. Then if $(\alpha_0, \alpha_1, \cdots, \alpha_l)$ is a fundamental system of positive roots, we know (Corollary 2 to Lemma 13 of [5(e)]) that it contains exactly one noncompact root, which we may assume to be α_0. Let Λ_i $0 \leqq i \leqq l$ be the linear functions on $\mathfrak{h}$ given by $\Lambda_i(H_{\alpha_j}) = \delta_{ij}$ $0 \leqq i, j \leqq l$ and let σ_i be an irreducible complex representation of G_c (see [5(f), § 6]) on a finite-dimensional Hilbert space V_i with the highest weight Λ_i. We assume that σ_i is unitary on V. Let ϕ_i be a unit vector in V_i belonging to the weight Λ_i. Then if $\psi_i(z) = (\phi_i, \sigma_i(z)\phi_i)$ $(z \varepsilon G_c)$, it follows from Lemmas 6 and 14 of [5(f)] that $\psi_i(\bar{x}) = \psi_{\Lambda_i}(x)$ $(0 \leqq i \leqq l)$ and

$$\psi_0(\bar{x}) = \psi_{\Lambda_0}(x) = \exp(\Lambda_0(\Gamma(x))) \qquad (x \varepsilon G).$$

Hence in particular $\psi_{\Lambda_0}(x)$ is never zero. Now if Λ is any linear function on $\mathfrak{h}$, it is obvious that $\Lambda = \lambda_0\Lambda_0 + \cdots + \lambda_l\Lambda_l$ where $\lambda_i = \Lambda(H_{\alpha_i})$ $0 \leqq i \leqq l$. Therefore if $\Lambda \varepsilon \mathfrak{F}_G(P)$, $\lambda_1, \cdots, \lambda_l$ are all non-negative integers while λ_0 is a *negative* real number. Moreover it follows again from Lemma 6 of [5(f)] that

$$\psi_\Lambda(x) = e^{\lambda_0\Lambda_0(\Gamma(x))}\psi_1^{\lambda_1}(\bar{x}) \cdots \psi_l^{\lambda_l}(\bar{x}) \qquad (x \varepsilon G).$$

In particular if λ_0 is also an integer, $\psi_\Lambda(x)$ depends only on $\bar{x}$ and so if we regard ψ_Λ as a function on G_0, we have

$$\psi_\Lambda = \psi_0^{\lambda_0}\psi_1^{\lambda_1} \cdots \psi_l^{\lambda_l}.$$

On the other hand if $m_0, m_1, \cdots, m_l$ are non-negative integers,

$$\psi = \psi_0^{m_0}\psi_1^{m_1} \cdots \psi_l^{m_l}$$

is a holomorphic function on G_c and again we can conclude from Lemma 6 of [5(f)] that

$$\psi(z) = (\phi, \sigma(z)\phi) \qquad (z \varepsilon G_c)$$

where σ is the irreducible complex representation of G_c on a finite-dimensional Hilbert space V with the highest weight $m_0\Lambda_0 + \cdots + m_l\Lambda_l$ and ϕ is a unit vector in V belonging to this weight. (We assume of course that σ is unitary on U). Then $\psi(u)$ $(u \varepsilon U)$ is a matrix coefficient of an irreducible representation of U and therefore the space spanned by all the right translates of ψ under U is irreducible under the corresponding representation of U. Thus the similarity between this case and that of ψ_Λ, $\mathfrak{H}_\Lambda$ and π_Λ discussed above is now quite obvious.

12. Proof of Lemma 22. In order to prove Lemma 22, we shall use a method which has been extensively used before by Weyl [11(a)] and Cartan [2(a)]. It is no longer necessary to assume that $\mathfrak{h}_0$ is maximal abelian in $\mathfrak{g}_0$, since this assumption plays no role whatever in our proof. However we still consider $\mathfrak{g}$ as a Hilbert space under the norm $\| X \|^2 = -B(\bar{\theta}(X), X)$ $(X \in \mathfrak{g})$ and denote by G_c a simply connected complex Lie group with the Lie algebra $\mathfrak{g}$. As before G_0, K_0 and U are the (real) analytic subgroups of G_c corresponding to $\mathfrak{g}_0$, $\mathfrak{k}_0$ and $\mathfrak{u} = \mathfrak{k}_0 + (-1)^{\frac{1}{2}}\mathfrak{p}_0$. We assume that $\int_{K_0} dk = 1$.

Let us introduce the following notation for the sake of convenience. Suppose μ and μ' are positive measures on two locally compact spaces E and E' respectively and f is an open continuous mapping of E into E' which is locally one-one on E. Then we write $d\mu \underset{f}{\sim} d\mu'$ if there exists a positive constant c with the following property. Let U be an open set in E such that f is univalent on U. Then $\mu(U) = c\mu'(f(U))$. If it is clear from the context which mapping f we have in mind, we write simply $d\mu \sim d\mu'$.

LEMMA [11] 31. $(k, X) \to k \exp X$ $(k \in K_0, X \in \mathfrak{p}_0)$ *is a one-one regular analytic mapping of* $K_0 \times \mathfrak{p}_0$ *onto* G_0.

Let ϕ denote this mapping. It is known (Cartan [2(b)], Mostow [9]) that ϕ is one-one and onto G_0. Also it is obviously analytic. Let f be a function on G_0 which is defined and analytic around $x = k \exp X$. Then if $Y \in \mathfrak{p}_0$ and $Z \in \mathfrak{k}_0$,

$$\left\{ \frac{d}{dt} f(k \exp(X + tY)) \right\}_{t=0} = (Y'f)(x)$$

$$\left\{ \frac{d}{dt} f(k \exp tZ \exp X) \right\}_{t=0} = (Z'f)(x) \qquad (t \in R)$$

where (see Chevalley [3, p. 157])

$$Y' = \{(1 - \exp(-adX))/adX\} Y$$

$$Z' = \exp(-adX) Z$$

and $(1 - \exp(-adX))/adX$ stands for the sum of the convergent series

$$\sum_{m \geqq 0}^{\infty} (-1)^m (adX)^m/(m+1)!.$$

[11] The proofs of Lemmas 31 and 33, which we present here, are substantially the same as those of Cartan. However in view of later applications, we are interested not only in verifying that certain determinants are different from zero but also in computing their actual value. This compels us to reproduce the whole argument.

Let T be the linear transformation of $\mathfrak{g}$ such that $TY = Y'$ and $TZ = Z'$. Since $(adX)\mathfrak{p} \subset \mathfrak{k}$ and $(adX)^2\mathfrak{p} \subset \mathfrak{p}$, it follows that

$$\exp(adX)Z' = Z$$

$$\exp(adX)Y' \equiv \{\sinh adX/adX\}Y \bmod k$$

where $(\sinh adX/adX) = s(X)$ is defined by the power series

$$\sum_{m\geqq 0} (adX)^{2m}/(2m+1)!.$$

Since only even powers of adX appear in this series $s(X)$ leaves $\mathfrak{p}$ invariant. Also $\bar{\theta}(X) = -X$ and therefore adX is self-adjoint. Hence $s(X)$ is a positive definite self-adjoint transformation. Therefore the same holds for its restriction $(s(X))_{\mathfrak{p}}$ on $\mathfrak{p}$. On the other hand it is obvious from the above congruence that

$$\det(\exp(adX)T) = \det((s(X))_{\mathfrak{p}}).$$

Since G_c is semisimple $\det(\exp(adX)) = 1$ and therefore

$$\det T = \det(s(X))_{\mathfrak{p}}.$$

The right side is positive since $(s(X))_{\mathfrak{p}}$ is positive definite. This proves that $\det T \neq 0$ and therefore ϕ is regular at (k, X).

Let dX denote the element of volume in p_0 corresponding to the Euclidean metric $\| X \|$ $(X \varepsilon \mathfrak{p}_0)$.

COROLLARY. $dx \sim \det(\sinh adX/adX)_{\mathfrak{p}}\, dkdX$ $(x = k \exp X)$.

This follows immediately from our calculation above.

We now state a lemma which will be needed frequently.

LEMMA 32. *Let M and N be two manifolds of class C^1 satisfying the countability axioms. Then if f is a differentiable mapping of M into N, and Q is a subset of M then $\partial\mathrm{im}\, f(Q) \leqq \partial\mathrm{im}\, Q$.*

Here we use the Brouwer-Urysohn-Menger definition [12] of dimension (Hurewicz and Wallman [6]). Although this result must be regarded as known, no easily accessible published proof of it seems to be available. Therefor we shall give a short proof in the Appendix (Section 13).

Let $\mathfrak{a}_{\mathfrak{p}_0}$ be a maximal abelian subspace of $\mathfrak{p}_0$. Then we have the following result [11] due to Cartan [2(a), p. 354].

[12] In order to distinguish it from the vector dimension of a complex vector space, we denote the topological dimension by "∂im."

LEMMA 33. $\mathfrak{p}_0 = \bigcup_{k \varepsilon K_0} \mathrm{Ad}(k)\mathfrak{a}_{\mathfrak{p}_0}$.

Extend $\mathfrak{a}_{\mathfrak{p}_0}$ to a maximal abelian subalgebra $\mathfrak{a}_0$ of $\mathfrak{g}_0$ and let $\mathfrak{a}_{\mathfrak{p}}$ and $\mathfrak{a}$ be the complexifications of $\mathfrak{a}_{\mathfrak{p}_0}$ and $\mathfrak{a}_0$ respectively in $\mathfrak{g}$. Then $\mathfrak{a}$ is a Cartan subalgebra of $\mathfrak{g}$. We consider the set Σ of all roots of $\mathfrak{g}$ (with respect to $\mathfrak{a}$) which do not vanish identically on $\mathfrak{a}_{\mathfrak{p}_0}$. For each $\alpha \varepsilon \Sigma$ let σ_α denote the hyperplane consisting of all points $H \varepsilon \mathfrak{a}_{\mathfrak{p}_0}$ such that $\alpha(H) = 0$ and let $\mathfrak{a}'_{\mathfrak{p}_0}$ denote the complement of $\bigcup_{\alpha \varepsilon \Sigma} \sigma_\alpha$ in $\mathfrak{a}_{\mathfrak{p}_0}$. Moreover let M be the set of all $k \varepsilon K_0$ such that $\mathrm{Ad}(k)H = H$ for every $H \varepsilon \mathfrak{a}_{\mathfrak{p}_0}$. Obviously M is a closed subgroup of K_0. Let $k \to \bar{k}$ $(k \varepsilon K_0)$ denote the natural mapping of K_0 on the factor space $\bar{K}_0 = K_0/M$ consisting of the cosets of the form kM. Consider the mapping ϕ of $\bar{K}_0 \times \mathfrak{a}_{\mathfrak{p}_0}$ into $\mathfrak{p}_0$ given by $\phi(\bar{k}, H) = H^{\bar{k}} = \mathrm{Ad}(k)H$ where k is any element in the coset $\bar{k}$. ϕ is obviously analytic. Let $\mathfrak{m}_0$ be the centralizer of $\mathfrak{a}_{\mathfrak{p}_0}$ in $\mathfrak{k}_0$, so that $\mathfrak{m}_0$ is the Lie algebra of M. Let f be a function on $\mathfrak{p}_0$ defined and analytic around $X = H_0^{\bar{k}_0}$ $(H_0 \varepsilon \mathfrak{a}_{\mathfrak{p}_0}, k_0 \varepsilon K_0)$. Then if $Z \varepsilon \mathfrak{k}_0$ and $H \varepsilon \mathfrak{a}_{\mathfrak{p}_0}$,

$$\left\{ \frac{d}{dt} f(\phi(\overline{k_0 \exp tZ}, H_0)) \right\}_{t=0} = \left\{ \frac{d}{dt} f(X - t\,\mathrm{Ad}(k_0)[H_0, Z]) \right\}_{t=0}$$

$$\left\{ \frac{d}{dt} f(\phi(\bar{k}_0, H_0 + tH)) \right\}_{t=0} = \left\{ \frac{d}{dt} f(X + t\,\mathrm{Ad}(k_0)H) \right\}_{t=0}$$

$(t \varepsilon R)$. Now if $H_0 \varepsilon \mathfrak{a}'_{\mathfrak{p}_0}$, $\dim([H_0, \mathfrak{k}]) = \dim \mathfrak{k} - \dim \mathfrak{m}$ (where $\mathfrak{m}$ is the complexification of $\mathfrak{m}_0$). Also $\mathfrak{a}_{\mathfrak{p}}$ and $(\mathrm{ad}\, H_0)\mathfrak{k}$ are mutually orthogonal. Hence

$$\dim(\mathfrak{a}_{\mathfrak{p}} + (\mathrm{ad}\, H_0)\mathfrak{k}) = \dim \mathfrak{a}_{\mathfrak{p}} + \dim \mathfrak{k} - \dim \mathfrak{m} = \dim \mathfrak{p}$$

from Lemma 4 of [5(b)]. This proves that

$$\mathrm{Ad}(k_0)(\mathfrak{a}_{\mathfrak{p}} + (\mathrm{ad}\, H_0)\mathfrak{k}) = \mathfrak{p}.$$

Since $\dim(\bar{K}_0 \times \mathfrak{a}_{\mathfrak{p}_0}) = \dim \mathfrak{k} - \dim \mathfrak{m} + \dim \mathfrak{a}_{\mathfrak{p}} = \dim \mathfrak{p} = \dim_R \mathfrak{p}_0$, it follows that ϕ is regular on $\bar{K}_0 \times \mathfrak{a}'_{\mathfrak{p}_0}$. Let $d\bar{k}$ denote the invariant measure on $\bar{K}_0$ such that $\int_{\bar{K}_0} d\bar{k} = 1$. Similarly let dH denote the element of volume in $\mathfrak{a}_{\mathfrak{p}_0}$ corresponding to the Euclidean metric $\| H \|$. Now introduce some (fixed) lexicographic order among roots and let Σ^+ be the set of positive roots in Σ under this order. For each root α select an element $X_\alpha \neq 0$ in $\mathfrak{g}$ such that $[H, X_\alpha] = \alpha(H)X_\alpha$ for all $H \varepsilon \mathfrak{a}$. Let $\mathfrak{n}^+$ and $\mathfrak{n}^-$ respectively be the subspaces of $\mathfrak{g}$ spanned by X_α and $X_{-\alpha}$ $(\alpha \varepsilon \Sigma^+)$. Then (see the proof of Lemma 4 of [5(b)])

$$\mathfrak{g} = \mathfrak{n}^- + \mathfrak{a}_\mathfrak{p} + \mathfrak{m} + \mathfrak{n}^+ = \mathfrak{k} + \mathfrak{a}_\mathfrak{p} + \mathfrak{n}^+ = \mathfrak{p} + \mathfrak{m} + \mathfrak{n}^+$$

where the sums are all direct. Hence

$$\mathfrak{k}/\mathfrak{m} \cong \mathfrak{g}/(\mathfrak{m} + \mathfrak{a}_\mathfrak{p} + \mathfrak{n}^+) \cong \mathfrak{p}/\mathfrak{a}_\mathfrak{p}$$

and

$$\mathfrak{g}/(\mathfrak{m} + \mathfrak{a}_\mathfrak{p} + \mathfrak{n}^+) \cong \mathfrak{n}^-.$$

Therefore we may identify $\mathfrak{k}/\mathfrak{m}$ and $\mathfrak{p}/\mathfrak{a}_\mathfrak{p}$ with $\mathfrak{n}^-$ in a natural way. Now if $H \varepsilon \mathfrak{a}_\mathfrak{p}$, $\mathfrak{m}$ is contained in the kernel of the mapping $Z \to [H, Z]$ $(Z \varepsilon \mathfrak{k})$ and therefore by going to the residue classes we get a mapping of $\mathfrak{k}/\mathfrak{m}$ into $\mathfrak{p}/\mathfrak{a}_\mathfrak{p}$. In view of the above identification, this gives a linear transformation T_H in $\mathfrak{n}^-$. It is clear that T_H coincides with the restriction of $\operatorname{ad} H$ on $\mathfrak{n}^-$. Therefore in view of our calculation above we can conclude that

$$dX \sim |\det T_H|\, dH d\bar{k} = \prod_{\alpha \varepsilon \Sigma^+} |\alpha(H)|\, dH d\bar{k} \qquad (X = H^{\bar{k}})$$

$(H \varepsilon \mathfrak{a}'_{\mathfrak{p}_0}, \bar{k} \varepsilon \bar{K}_0)$. We shall need this result a little later.

Let α be a root in Σ. Since α takes real values on $\mathfrak{a}_{\mathfrak{p}_0}$, we can find an element $X \neq 0$ in $\mathfrak{g}_0$ such that $[H, X] = \alpha(H) X$ for all $H \varepsilon \mathfrak{a}_{\mathfrak{p}_0}$. We claim $X \notin \mathfrak{p}_0$. For otherwise since $[\mathfrak{p}_0, \mathfrak{p}_0] \subset \mathfrak{k}_0$, we get $[H, X] \varepsilon \mathfrak{k} \cap \mathfrak{p}_0 = 0$. This however is impossible since $X \neq 0$ and $\alpha(H) \neq 0$ for a suitable H in $\mathfrak{a}_{\mathfrak{p}_0}$. Similarly we prove that $X \notin \mathfrak{k}_0$. Therefore $X = Y + Z$ $(Y \varepsilon \mathfrak{p}, Z \varepsilon \mathfrak{k}_0)$ and $Y \neq 0$, $Z \neq 0$. Then comparing components in $\mathfrak{k}_0$ and $\mathfrak{p}_0$ of both sides of the equation $[H, X] = \alpha(H) X$, we find that $[H, Y] = \alpha(H) Z$ and $[H, Z] = \alpha(H) Y$. Thus we have obtained the following result (see Cartan [2(a), p. 357]).

LEMMA 34. *For each $\alpha \varepsilon \Sigma$ we can choose nonzero elements Y_α, Z_α in $\mathfrak{p}_0$ and $\mathfrak{k}_0$ respectively such that*

$$[H, Y_\alpha] = \alpha(H) Z_\alpha, \qquad [H, Z_\alpha] = \alpha(H) Y_\alpha$$

for all $H \varepsilon \mathfrak{a}_\mathfrak{p}$.

Let M_α $(\alpha \varepsilon \Sigma^+)$ denote the set of all elements $k \varepsilon K_0$ such that $\operatorname{Ad}(k) H = H$ for every $H \varepsilon \sigma_\alpha$. Then M_α is a compact subgroup of K_0 containing M. Moreover it is obvious that the element Z_α of the above lemma does not lie in $\mathfrak{m}_0$ while it is certainly contained in the Lie algebra of M_α. Hence $\dim M_\alpha > \dim M$. Let ϕ_α be the mapping of $(K_0/M_\alpha) \times \sigma_\alpha$ into $\mathfrak{p}_0$ defined by $\phi_\alpha(k^*, H) = \operatorname{Ad}(k) H$ where $k \varepsilon K_0$, $H \varepsilon \sigma_\alpha$ and $k^* = k M_\alpha$.

It is obvious that ϕ_α is an analytic mapping and therefore it follows from Lemma 32 that

$$\partial\text{im}\, \phi_\alpha((K_0/M_\alpha)\times\sigma_\alpha) \leqq \partial\text{im}((K_0/M_\alpha)\times\sigma_\alpha) \leqq \dim_R \mathfrak{p}_0 - 2$$

since $\dim(K_0/M_\alpha) < \dim(K_0/M)$ and

$$\partial\text{im}\, \sigma_\alpha = \dim_R \sigma_\alpha = \dim_R \mathfrak{a}_{\mathfrak{p}_0} - 1.$$

On the other hand it is clear that

$$\phi_\alpha((K_0/M_\alpha)\times\sigma_\alpha) = \phi(\bar{K}_0\times\sigma_\alpha)$$

and therefore

$$\phi((K_0/M)\times(\bigcup_{\alpha\varepsilon\Sigma^+}\sigma_\alpha)) = \bigcup_{\alpha\varepsilon\Sigma^+}\phi_\alpha((K_0/M_\alpha)\times\sigma_\alpha).$$

Hence if we denote this set by $\mathfrak{p}_s$,

$$\partial\text{im}\, \mathfrak{p}_s \leqq \partial\text{im}\, \mathfrak{p}_0 - 2.$$

Therefore the complement ${}^c\mathfrak{p}_s$ of $\mathfrak{p}_s$ in $\mathfrak{p}_0$ is connected (see Hurewicz and Wallman [6, p. 48, Theorem IV 4]). On the other hand let

$$\mathfrak{p}'_0 = \phi(\bar{K}_0\times\mathfrak{a}'_{\mathfrak{p}_0}).$$

Since ϕ is regular and therefore open on $\bar{K}_0\times\mathfrak{a}'_{\mathfrak{p}_0}$, $\mathfrak{p}'_0$ is open in $\mathfrak{p}_0$. Let $r_0 = \dim_R(\mathfrak{a}_{\mathfrak{p}_0} + \mathfrak{m}_0)$ and for any $X \varepsilon \mathfrak{p}_0$ let $r(X)$ denote the complex dimension of the centralizer of X in $\mathfrak{g}$. Then it is obvious that $r(X) = r_0$ if $X \varepsilon \mathfrak{p}'_0$ while $r(X) > r_0$ if $X \varepsilon \mathfrak{p}_s$. Hence $\mathfrak{p}' \subset {}^c\mathfrak{p}_s$. Finally since $\bar{K}_0$ is compact, it follows immediately that $\mathfrak{p}'_0$ is closed in ${}^c\mathfrak{p}_s$. Therefore since ${}^c\mathfrak{p}_s$ is connected and $\mathfrak{p}'_0$ is not empty (because $\mathfrak{a}'_{\mathfrak{p}_0} \subset \mathfrak{p}'_0$) we conclude that $\mathfrak{p}'_0 = {}^c\mathfrak{p}_s$. Now let λ be an indeterminate. Then it is clear that there exists a polynomial function F on $\mathfrak{p}$ such that if $X \varepsilon \mathfrak{p}$, $F(X)$ is the coefficient of λ^{r_0} in the characteristic polynomial of $\operatorname{ad} X$ (in λ). It is obvious that $r(X) > r_0$ if and only if $F(X) = 0$. Hence F is not identically zero and $\mathfrak{p}_s$ is contained in the set of zeros of F on $\mathfrak{p}_0$. This shows that ${}^c\mathfrak{p}_s = \mathfrak{p}'_0$ is everywhere dense in $\mathfrak{p}_0$. But $\bar{K}_0$ is compact, and $\| H^{\bar{k}} \| = \| H \|$ ($\bar{k} \varepsilon \bar{K}$, $H \varepsilon \mathfrak{a}_{\mathfrak{p}_0}$) and so it follows that $\phi(\bar{K}_0\times\mathfrak{a}_{\mathfrak{p}_0})$ is closed in $\mathfrak{p}_0$. Therefore $\phi(\bar{K}_0\times\mathfrak{a}_{\mathfrak{p}_0}) = \mathfrak{p}_0$. This completes the proof of Lemma 33.

Let M' be the normalizer of $\mathfrak{a}_{\mathfrak{p}_0}$ in K. Then M' is a closed subgroup of K_0 and M is a normal subgroup of M'. If $H \varepsilon \mathfrak{a}_{\mathfrak{p}_0}$, $\operatorname{ad} H$ is self-adjoint and therefore $(\operatorname{ad} H)^2 X = 0$ implies $(\operatorname{ad} H)X = 0$ ($X \varepsilon \mathfrak{g}$). From this it follows immediately that the Lie algebra of M' is also $\mathfrak{m}_0$ and therefore $W = M'/M$

is discrete. But since it is also compact, W must be finite. It operates as a group of linear transformations on $\mathfrak{a}_{\mathfrak{p}_0}$ as follows:

$$sH = \mathrm{Ad}(k)H \qquad (s \in W)$$

where k is any element in M' belonging to the coset s.

LEMMA 35. *Let H be an element in $\mathfrak{a}'_{\mathfrak{p}_0}$. Then if $k \in K_0$ and $\mathrm{Ad}(k)H \in \mathfrak{a}_{\mathfrak{p}_0}$, k must lie in M'.*

Since $H \in \mathfrak{a}'_{\mathfrak{p}_0}$, $\mathfrak{a}_{\mathfrak{p}_0}$ is exactly the centralizer of H in $\mathfrak{p}_0$. Therefore $\mathrm{Ad}(k)\mathfrak{a}_{\mathfrak{p}_0}$ is the centralizer of $\mathrm{Ad}(k)H$ in $\mathfrak{p}_0$. But since $\mathrm{Ad}(k)H \in \mathfrak{a}_{\mathfrak{p}_0}$ and $\mathfrak{a}_{\mathfrak{p}_0}$ is abelian, $\mathfrak{a}_{\mathfrak{p}_0} \subset \mathrm{Ad}(k)\mathfrak{a}_{\mathfrak{p}_0}$. However $\mathfrak{a}_{\mathfrak{p}_0}$ and $\mathrm{Ad}(k)\mathfrak{a}_{\mathfrak{p}_0}$ obviously have the same dimension over R. Therefore $\mathfrak{a}_{\mathfrak{p}_0} = \mathrm{Ad}(k)\mathfrak{a}_{\mathfrak{p}_0}$ and so $k \in M'$.

W also operates on $\bar{K}_0$ on the right as follows. Let $\bar{k} \in \bar{K}_0$ and $s \in W$. Since M is a normal subgroup of M', $\bar{k}s = km'M$ where k and m' are any two elements of K_0 and M' respectively lying in the cosets $\bar{k}$ and s. It is obvious that $\phi(\bar{k}s, s^{-1}H) = \phi(\bar{k}, H)$ $(H \in \mathfrak{a}_{\mathfrak{p}_0})$. Conversely if $H \in \mathfrak{a}'_{\mathfrak{p}_0}$ and $\phi(\bar{k}, H) = \phi(\bar{k}_1, H_1)$ $(k, k_1 \in K_0, H_1 \in \mathfrak{a}_{\mathfrak{p}_0})$, it follows from Lemma 35 that $k_1^{-1}k \in M'$. This shows that $\bar{k}_1 = \bar{k}s$ and $H_1 = s^{-1}H$ for some $s \in W$. Then if w_0 is the order of the group W, it follows that every element in $\mathfrak{p}'_0$ has exactly w_0 distinct pre-images in $\bar{K}_0 \times \mathfrak{a}'_{\mathfrak{p}_0}$ under ϕ. We have seen above that there exists a polynomial function $F \neq 0$ on $\mathfrak{p}_0$ such that $F(X) = 0$ $(X \in \mathfrak{p}_0)$ if and only if $X \notin \mathfrak{p}'_0$. This shows that the complement of $\mathfrak{p}'_0$ in $\mathfrak{p}_0$ is of Euclidean measure zero. Similarly it is obvious that the complement of $\mathfrak{a}'_{\mathfrak{p}_0}$ in $\mathfrak{a}_{\mathfrak{p}}$ is of Euclidean measure zero. Moreover we have seen above that

$$dX \sim \prod_{\alpha \in \Sigma^+} |\alpha(H)|\, dH d\bar{k} \, (X = H^{\bar{k}}, H \in \mathfrak{a}'_{\mathfrak{p}_0}, \bar{k} \in \bar{K}_0),$$

and so we have the following result.

LEMMA 36. *There exists a positive constant c with the following property. If f is any continuous function on $\mathfrak{p}_0$ vanishing outside a compact set,*

$$\int_{\mathfrak{p}_0} f(X)dX = c\int_{\bar{K}_0 \times \mathfrak{A}_{\mathfrak{p}_0}} f(H^{\bar{k}}) \prod_{\alpha \in \Sigma^+} |\alpha(H)|\, d\bar{k}dH.$$

On the other hand we have the following lemma.

LEMMA 37. *Let $\mathfrak{b}_0$ be a connected component of $\mathfrak{a}'_{\mathfrak{p}_0}$. Then*

$$\mathfrak{a}'_{\mathfrak{p}_0} = \bigcup_{s \in W} s\mathfrak{b}_0.$$

Let $\bar{\mathfrak{b}}_0$ denote the closure of $\mathfrak{b}_0$ in $\mathfrak{a}_{\mathfrak{p}_0}$. Since $\bar{\mathfrak{b}}_0$ is obviously closed

in $\mathfrak{a}'_{\mathfrak{p}_0}$, $\bar{\mathfrak{b}}_0 \cap \mathfrak{a}'_{\mathfrak{p}_0} = \mathfrak{b}_0$. Consider $\phi(\bar{K}_0 \times \mathfrak{b}_0) \subset \mathfrak{p}'_0$. Since ϕ is regular on $\bar{K}_0 \times \mathfrak{a}'_{\mathfrak{p}_0}$, $\phi(\bar{K}_0 \times \mathfrak{b}_0)$ is open. Moreover since $\bar{K}_0$ is compact it follows easily that $\phi(\bar{K}_0 \times \bar{\mathfrak{b}}_0)$ is closed in $\mathfrak{p}_0$. Hence $\phi(\bar{K}_0 \times \mathfrak{b}_0) = \phi(\bar{K}_0 \times \bar{\mathfrak{b}}_0) \cap \mathfrak{p}'_0$ is closed in $\mathfrak{p}'_0$. But we know that $\mathfrak{p}'_0$ is connected and therefore $\phi(\bar{K}_0 \times \mathfrak{b}_0) = \mathfrak{p}'_0$. Hence in particular if $H \varepsilon \mathfrak{a}'_{\mathfrak{p}_0}$, $H = \mathrm{Ad}(k)H_0$ for some $k \varepsilon K_0$ and $H_0 \varepsilon \mathfrak{b}_0$. But then it follows from Lemma 35 that $k \varepsilon M'$ and therefore $H = sH_0$ for some $s \varepsilon W$. This proves that $\mathfrak{a}'_{\mathfrak{p}_0} \subset \bigcup_{s \varepsilon W} s\mathfrak{b}_0$. Since the reverse inclusion is obvious we get the lemma.

COROLLARY. *$\mathfrak{a}'_{\mathfrak{p}_0}$ has only a finite number of connected components.*

Now as we have seen earlier, $\prod_{\alpha \varepsilon \Sigma^+} |\alpha(H)|^2$ $(H \varepsilon \mathfrak{a}_{\mathfrak{p}_0})$ is the determinant of the linear transformation in $\mathfrak{p}/\mathfrak{a}_{\mathfrak{p}}$ corresponding to $(adH)^2$. From this it follows that the value of this expression does not change if we replace H by sH $(s \varepsilon W)$. Therefore it is obvious from Lemma 37 that

$$\int_{K_0 \times \mathfrak{A}_{\mathfrak{p}_0}} f(\mathrm{Ad}(k)H) \prod_{\alpha \varepsilon \Sigma^+} |\alpha(H)|\, dkdH = w \int_{K_0 \times \mathfrak{b}_0} f(\mathrm{Ad}(k)H) \prod_{\alpha \varepsilon \Sigma^+} |\alpha(H)|\, dkdH$$

where w is the number of connected components of $\mathfrak{a}'_{\mathfrak{p}_0}$, $\mathfrak{b}_0$ is any such component and f is a continuous function on $\mathfrak{p}_0$ vanishing outside a compact set. Combining this with Lemma 31 and its corollary, we obtain the following result.

LEMMA 38. *It is possible to normalize the Haar measure dx on G_0 in such a way that the following condition is fulfilled. If f is a continuous function on G_0 vanishing outside a compact set and $\mathfrak{b}_0$ is any connected component of $\mathfrak{a}'_{\mathfrak{p}_0}$,*

$$\int_{G_0} f(x)\,dx = \int_{K_0 \times \mathfrak{p}_0} f(k \exp X) \det (\sinh \mathrm{ad}\, X/\mathrm{ad}\, X)_{\mathfrak{p}}\, dkdX$$
$$= cw \int_{\mathfrak{b}_0} \prod_{\alpha \varepsilon \Sigma^+} |e^{\alpha(H)} - e^{-\alpha(H)}|\, dH \int_{K_0 \times K_0} f(k \exp H\, k')\, dkdk'$$

where w is the number of connected components of $\mathfrak{a}'_{\mathfrak{p}_0}$ and c is a positive constant given by the relation

$$\int_{\mathfrak{p}_0} e^{-\|X\|^2} \det(\sinh adX/adX)_{\mathfrak{p}}\, dX = c \int_{\mathfrak{A}_{\mathfrak{p}_0}} e^{-\|H\|^2} \prod_{\alpha \varepsilon \Sigma^+} |e^{\alpha(H)} - e^{-\alpha(H)}|\, dH.$$

For the proof we have only to notice the fact that

$$\det(\sinh adH/adH) = \prod_{\alpha \varepsilon \Sigma^+} (\sinh \alpha(H)/\alpha(H)) \qquad (H \varepsilon \mathfrak{a}_{\mathfrak{p}_0}).$$

The following lemma is also due to Cartan [2(a)].

LEMMA 39. *Let $'\mathfrak{a}_{\mathfrak{p}_0}$ be any maximal abelian subspace of $\mathfrak{p}_0$. Then*

$$'\mathfrak{a}_{\mathfrak{p}_0} = \mathrm{Ad}(k)\mathfrak{a}_{\mathfrak{p}_0}$$

for some $k \varepsilon K_0$.

Define $\mathfrak{a}'_{\mathfrak{p}_0}$ as above and choose $H_0 \varepsilon \mathfrak{a}'_{\mathfrak{p}_0}$. Then $\mathfrak{a}_{\mathfrak{p}_0}$ is exactly the centralizer of H_0 in $\mathfrak{p}_0$. If we apply Lemma 33 to $'\mathfrak{a}_{\mathfrak{p}_0}$ (instead of $\mathfrak{a}_{\mathfrak{p}_0}$), it follows that $H_0 \varepsilon \mathrm{Ad}(k^{-1})'\mathfrak{a}_{\mathfrak{p}_0}$ for some $k \varepsilon K_0$. Since $'\mathfrak{a}_{\mathfrak{p}_0}$ is abelian the same holds for $\mathrm{Ad}(k^{-1})'\mathfrak{a}_{\mathfrak{p}_0}$ and therefore $\mathrm{Ad}(k^{-1})'\mathfrak{a}_{\mathfrak{p}_0} \subset \mathfrak{a}_{\mathfrak{p}_0}$ or $'\mathfrak{a}_{\mathfrak{p}_0} \subset \mathrm{Ad}(k)\mathfrak{a}_{\mathfrak{p}_0}$. But then since $'\mathfrak{a}_{\mathfrak{p}_0}$ is maximal abelian in $\mathfrak{p}_0$ it is obvious that $'\mathfrak{a}_{\mathfrak{p}_0} = \mathrm{Ad}(k)\mathfrak{a}_{\mathfrak{p}_0}$.

Let $\mathfrak{A}$ be the analytic subgroup of G_0 corresponding to $\mathfrak{a}_{\mathfrak{p}_0}$. We conclude from Lemma 31 that $\mathfrak{A}$ is closed and $H \to \exp H$ $(H \varepsilon \mathfrak{a}_{\mathfrak{p}_0})$ is a topological mapping of $\mathfrak{a}_{\mathfrak{p}_0}$ onto $\mathfrak{A}$. Moreover as we have seen in Section 8,

$$\begin{aligned} |D(\exp H)| &= \prod_{\alpha \varepsilon \Sigma} |e^{\alpha(H)} - e^{-\alpha(H)}| \\ &= \prod_{\alpha \varepsilon \Sigma^+} |e^{\alpha(H)} - e^{-\alpha(H)}|^2 \qquad (H \varepsilon \mathfrak{a}_{\mathfrak{p}_0}). \end{aligned}$$

Therefore if we define $\mathfrak{A}'$ as in Lemma 22, it is clear that $\mathfrak{A}' = \exp(\mathfrak{a}'_{\mathfrak{p}_0})$ and the number of connected component of $\mathfrak{A}'$ and $\mathfrak{a}'_{\mathfrak{p}_0}$ is the same. Moreover it follows from Lemma 39 without difficulty that the numbers c and w of Lemma 38 are independent of the choice of $\mathfrak{a}_{\mathfrak{p}_0}$. Therefore the second statement of Lemma 22 is now obvious.

It remains to prove the first part of Lemma 22. Let

$$e(X) = \exp(-1)^{\frac{1}{2}}X \varepsilon U \qquad (X \varepsilon \mathfrak{p}_0).$$

LEMMA 40. *$\mathfrak{A}^* = e(\mathfrak{a}_{\mathfrak{p}_0})$ is compact.*

Since U is compact it is enough to prove that $\mathfrak{A}^*$ is closed in U. Let $\bar{\mathfrak{A}}^*$ denote the closure of $\mathfrak{A}^*$ in U. Obviously $\mathfrak{A}^*$ and therefore $\bar{\mathfrak{A}}^*$ is a connected Lie subgroup of U. G_c being simply connected, we "extend" θ to a (complex) automorphism of G_c so that $\theta(\exp X) = \exp \theta(X)$ $(X \varepsilon \mathfrak{g})$. Since $\theta(H) = -H$ $(H \varepsilon \mathfrak{a}_{\mathfrak{p}_0})$, it is clear that $\theta(a) = a^{-1}$ for $a \varepsilon \mathfrak{A}^*$ and therefore by continuity also for $a \varepsilon \bar{\mathfrak{A}}^*$. Hence if X is an element in $\mathfrak{u}$ which lies in the Lie algebra of $\bar{\mathfrak{A}}^*$, $\theta(X) = -X$ and therefore $X \varepsilon \mathfrak{u} \cap \mathfrak{p} = (-1)^{\frac{1}{2}}\mathfrak{p}_0$. But since $\mathfrak{a}_{\mathfrak{p}_0}$ is maximal abelian in $\mathfrak{p}_0$, it is obvious that $X \varepsilon (-1)^{\frac{1}{2}}\mathfrak{a}_{\mathfrak{p}_0}$. This proves that $\mathfrak{A}^*$ and $\bar{\mathfrak{A}}^*$ have the same Lie algebra and therefore $\mathfrak{A}^* = \bar{\mathfrak{A}}^*$.

Define $\mathfrak{A}^{*\prime}$ as in Lemma 22. Then $e(H) \varepsilon \mathfrak{A}^{*\prime}$ $(H \varepsilon \mathfrak{a}_{\mathfrak{p}_0})$ if and only if

$$\prod_{\alpha \varepsilon \Sigma^+} |\sin \alpha(H)| \neq 0.$$

Put $U_* = U/K_0$ and $u_* = uK_0$ $(u \varepsilon U)$. We define an analytic mapping ψ of $\bar{K}_0 \times \mathfrak{A}^*$ into U_* as follows:

$$\psi(\bar{k}, a) = (ka)_* \qquad (\bar{k} \varepsilon \bar{K}, a \varepsilon \mathfrak{A}^*)$$

where k is any element in the coset $\bar{k}$.

LEMMA 41. *ψ is regular and open on $\bar{K}_0 \times \mathfrak{A}^{*\prime}$ and $\psi(\bar{K} \times \mathfrak{A}^*) = U_*$.*

Let u be an element in U and f a function which is defined and analytic on some neighborhood of u in U. Suppose $u = ka$ where $k \varepsilon K_0$ and $a \varepsilon \mathfrak{A}^*$. Then if $H \varepsilon (-1)^{\frac{1}{2}}\mathfrak{a}_{\mathfrak{p}_0}$ and $X \varepsilon k_0$,

$$\left\{ \frac{d}{dt} f(ka \exp tH) \right\}_{t=0} = (Hf)(u)$$

$$\left\{ \frac{d}{dt} f(k \exp tX\, a) \right\}_{t=0} = (X'f)(u)$$

where $X' = \mathrm{Ad}(a^{-1})X$. Let T denote the linear mapping of $\mathfrak{k} + \mathfrak{a}_{\mathfrak{p}}$ into $\mathfrak{g}$ given by

$$H \to H \qquad (H \varepsilon \mathfrak{a}_{\mathfrak{p}})$$

$$X \to \mathrm{Ad}(a^{-1})X \qquad (X \varepsilon \mathfrak{k}).$$

Since $\mathfrak{m}$ lies in the kernel of T, we get a linear mapping $\bar{T}$ of $(\mathfrak{k} + \mathfrak{a}_{\mathfrak{p}})/\mathfrak{m}$ into $\mathfrak{g}/\mathfrak{k}$ by going over to the factor spaces. Now suppose $T(H + X) \varepsilon \mathfrak{k}$ $(H \varepsilon \mathfrak{a}_{\mathfrak{p}}, X \varepsilon \mathfrak{k})$. Then $H + \mathrm{Ad}(a^{-1})X \varepsilon \mathfrak{k}$. But

$$\mathrm{Ad}(a^{-1})X \equiv \tfrac{1}{2}(\mathrm{Ad}(a^{-1}) - \mathrm{Ad}(a))X \bmod \mathfrak{k}$$

and $(\mathrm{Ad}(a^{-1}) - \mathrm{Ad}(a))X \varepsilon \mathfrak{p}$. Hence

$$H + \tfrac{1}{2}(\mathrm{Ad}(a^{-1}) - \mathrm{Ad}(a))X \varepsilon \mathfrak{k} \cap \mathfrak{p} = 0.$$

However it is easy to check that $(\mathrm{Ad}(a^{-1}) - \mathrm{Ad}(a))\mathfrak{g}$ and $\mathfrak{a}_{\mathfrak{p}}$ are orthogonal. Therefore $H = 0$ and $\frac{1}{2}(Ad(a^{-1}) - Ad(a))X = 0$. Now if in particular $a \varepsilon \mathfrak{A}^{*\prime}$, this clearly implies that $X \varepsilon \mathfrak{m}$. This shows that if $a \varepsilon \mathfrak{A}^{*\prime}$, the kernel of $\bar{T}$ is zero. Moreover as we have seen before $\mathfrak{g} = \mathfrak{k} + \mathfrak{a}_{\mathfrak{p}} + \mathfrak{n}^+ = \mathfrak{p} + \mathfrak{m} + \mathfrak{n}^+$ and therefore

$$(\mathfrak{k} + \mathfrak{a}_{\mathfrak{p}})/\mathfrak{m} \cong \mathfrak{g}/(\mathfrak{m} + \mathfrak{n}^+) \cong \mathfrak{p} \cong \mathfrak{g}/\mathfrak{k}$$

Hence $(\mathfrak{k} + \mathfrak{a}_{\mathfrak{p}})/\mathfrak{m}$ and $\mathfrak{g}/\mathfrak{k}$ have the same dimension. This proves that $\bar{T}$ is an isomorphism of $(\mathfrak{k} + \mathfrak{a}_{\mathfrak{p}})/\mathfrak{m}$ onto $\mathfrak{g}/\mathfrak{k}$ and therefore ψ is open and regular on $\bar{K}_0 \times \mathfrak{A}^{*\prime}$. Let du_* and $d\bar{k}$ denote the invariant measures on the

homogeneous spaces U_* and $\bar{K}_0$ respectively. We assume that

$$\int_{U_*} du_* = \int_{\bar{K}_0} d\bar{k} = 1.$$

Since $\mathfrak{g} = \mathfrak{n}^- + \mathfrak{a}_\mathfrak{p} + \mathfrak{m} + \mathfrak{n}^+$,

$$\mathfrak{g}/(\mathfrak{m} + \mathfrak{n}^+) \cong \mathfrak{a}_\mathfrak{p} + \mathfrak{n}^-$$

and therefore if we identify $(\mathfrak{k} + \mathfrak{a}_\mathfrak{p})/\mathfrak{m}$ and $\mathfrak{g}/\mathfrak{k}$ with $\mathfrak{a}_\mathfrak{p} + \mathfrak{n}^-$ under the isomorphism indicated above, $\bar{T}$ may be regarded as a linear mapping of $\mathfrak{a}_\mathfrak{p} + \mathfrak{n}^-$ into itself. If $Z \varepsilon \mathfrak{n}^-$,

$$\begin{aligned} \mathrm{Ad}(a^{-1})(Z + \theta(Z)) &\equiv \tfrac{1}{2}(\mathrm{Ad}(a^{-1}) - \mathrm{Ad}(a))(Z + \theta(Z)) \\ &\equiv Z' - \theta(Z') \bmod \mathfrak{k} \end{aligned}$$

where $Z' = \frac{1}{2}(\mathrm{Ad}(a^{-1}) - \mathrm{Ad}(a))Z$. From this it follows easily that

$$\bar{T}(H + Z) = H + \tfrac{1}{2}(\mathrm{Ad}(a^{-1}) - \mathrm{Ad}(a))Z \qquad (H \varepsilon \mathfrak{a}_\mathfrak{p}, Z \varepsilon \mathfrak{n}^-)$$

and therefore

$$|\det \bar{T}| = 2^{-q} |D(a)|^{\frac{1}{2}}$$

where q is the number of roots in Σ^+. This proves that

$$du_* \sim |D(a)|^{\frac{1}{2}} da d\bar{k} \qquad (u_* = \psi(\bar{k}, a))$$

if $\bar{k} \varepsilon \bar{K}_0$ and $a \varepsilon \mathfrak{A}^{*\prime}$ and da denotes the Haar measure on $\mathfrak{A}^*$. We shall need this result presently.

For each $\alpha \varepsilon \Sigma^+$ we consider the character ξ_α of $\mathfrak{A}^*$ given by $\xi_\alpha(e(H)) = \exp((-1)^{\frac{1}{2}}\alpha(H))$ $(H \varepsilon \mathfrak{a}_{\mathfrak{p}_0})$. Then obviously $a \to \xi_\alpha(a^2)$ $(a \varepsilon \mathfrak{A}^*)$ is also a non-trivial character of $\mathfrak{A}^*$. Hence its kernel $\mathfrak{A}^*_\alpha$ is a closed subgroup of $\mathfrak{A}^*$ and $\dim \mathfrak{A}^*_\alpha = \dim \mathfrak{A}^* - 1$. Let M_α be the subgroup of K_0 consisting of all $k \varepsilon K_0$ such that $(ka)_* = a_*$ for every $a \varepsilon \mathfrak{A}^*_\alpha$. Obviously M_α is closed and it contains M. Now consider the elements Y_α, Z_α of Lemma 34. If $H \varepsilon \mathfrak{a}_{\mathfrak{p}_0}$ it is obvious that

$$\mathrm{Ad}(e(H))Z_\alpha = \cos\alpha(H) Z_\alpha + (-1)^{\frac{1}{2}} \sin\alpha(H) Y_\alpha.$$

Moreover if $e(H) \varepsilon \mathfrak{A}^*_\alpha$, $\alpha(H)/\pi$ is an integer and therefore

$$\mathrm{Ad}(a) Z_\alpha = \pm Z_\alpha$$

for $a \varepsilon \mathfrak{A}^*_\alpha$. From this we conclude immediately that Z_α lies in the Lie algebra of M_α. Since $Z_\alpha \notin \mathfrak{m}_0$, $\dim M_\alpha > \dim M$. But $\psi(\bar{K}_0 \times \mathfrak{A}^*_\alpha)$ may be regarded as the image of $(K_0/M_\alpha) \times \mathfrak{A}^*_\alpha$ under an analytic mapping and therefore it follows from Lemma 32 that

$$\dim \psi(\bar{K}_0 \times \mathfrak{A}^*_\alpha) \leqq \dim K_0/M_\alpha + \dim \mathfrak{A}^*_\alpha$$
$$\leqq \dim \bar{K}_0 + \dim \mathfrak{A}^* - 2 = \dim U_* - 2.$$

Hence if $V = \bigcup_{\alpha \varepsilon \Sigma^+} \psi(\bar{K}_0 \times \mathfrak{A}^*_\alpha)$, V is compact and $\partial \mathrm{im}\, V \leqq \dim U_* - 2$ and therefore (see Hurewicz and Wallman [6, p. 48]) the complement ${}^cV$ of V in U_* is connected. Moreover ${}^cV$ is a dense subset of U_*. Now consider

$$U'_* = {}^cV \cap \psi(\bar{K}_0 \times \mathfrak{A}^*).$$

Since $\bar{K}_0 \times \mathfrak{A}^*$ is compact, U'_* is closed in ${}^cV$. On the other hand V and $\psi(\bar{K}_0 \times \mathfrak{A}^{*\prime})$ cannot have any point in common (see Corollary 1 to Lemma 42 below) and therefore

$$U'_* = \psi(\bar{K}_0/\mathfrak{A}^{*\prime}).$$

Since ψ is open on $\bar{K}_0 \times \mathfrak{A}^{*\prime}$, U'_* is a nonempty open subset of U_*. Therefore ${}^cV$ being connected, we can conclude that $U'_* = {}^cV$. This shows that the compact set $\psi(\bar{K}_0 \times \mathfrak{A}^*)$ is dense in U_*. Hence $\psi(\bar{K}_0 \times \mathfrak{A}^*) = U_*$ and Lemma 41 is proved.

For any $a \varepsilon \mathfrak{A}^*$ and $s \varepsilon W$, define $a^s = kak^{-1}$ where k is some element of M' lying in the coset s. Then $a^s \varepsilon \mathfrak{A}^*$ and $|D(a^s)| = |D(a)|$. Hence $(\mathfrak{A}^{*\prime})^s = \mathfrak{A}^{*\prime}$.

LEMMA 42. *Suppose* $(ka_1)_* = (a_2)_*$ $(k \varepsilon K_0; a_1, a_2 \varepsilon \mathfrak{A}^*)$. *Then if* $a_1 \varepsilon \mathfrak{A}^{*\prime}$, *$k$ lies in M' and a_2 in* $\mathfrak{A}^{*\prime}$.

For $ka_1 = a_2k'$ where $k' \varepsilon K_0$. Applying the automorphism θ to both sides we get $ka_1^{-1} = a_2^{-1}k'$ and therefore $ka_1^2k^{-1} = a_2^2$. Since $a_1 \varepsilon \mathfrak{A}^{*\prime}$, $\mathfrak{a}_{\mathfrak{p}_0}$ is exactly the set of all elements $H \varepsilon \mathfrak{p}_0$ such that $\mathrm{Ad}(a_1^2)H = H$. Therefore it follows from the above relation that $\mathrm{Ad}(k)\mathfrak{a}_{\mathfrak{p}_0} = \mathfrak{a}_{\mathfrak{p}_0}$ and so $k \varepsilon M'$. Also it is now clear that $a_2 \varepsilon \mathfrak{A}^{*\prime}$.

COROLLARY 1. *Suppose $a \varepsilon \mathfrak{A}^{*\prime}$ and $\bar{k} \varepsilon \bar{K}_0$. Then if w_0 is the order of W, $\psi(\bar{k}, a)$ has exactly w_0 distinct preimages in $\bar{K}_0 \times \mathfrak{A}^*$ namely $(\bar{k}s^{-1}, a^s)$ $(s \varepsilon W)$.*

Choose an element k in the coset $\bar{k}$. Then if $\psi(\bar{k}_1, a_1) = \psi(\bar{k}, a)$ $(k_1 \varepsilon K_0$, $a_1 \varepsilon \mathfrak{A}^*)$, it follows from Lemma 42 that $k^{-1}k_1 \varepsilon M'$ and therefore if s^{-1} denotes the corresponding element of W, $\bar{k}_1 = \bar{k}s^{-1}$ and $a_1 = a^s$.

COROLLARY 2. *There exists a positive number c with the following property. If f is any continuous function on U_*,*

$$\int_{U_*} f(u_*)\,du_* = c \int_{\mathfrak{A}^*} |D(a)|^{\frac{1}{2}}\,da \int_{\bar{K}_0} f(\psi(\bar{k}, a))\,d\bar{k}.$$

We have seen that $du_* \sim |D(a)|^{\frac{1}{2}} da d\bar{k}$ if $u_* = \psi(\bar{k}, a)$ and $\bar{k} \varepsilon \bar{K}_0$, $a \varepsilon \mathfrak{A}^{*\prime}$. Now if we define $V = \bigcup_{\alpha \varepsilon \Sigma^+} \psi(\bar{K}_0 \times \mathfrak{A}^*_\alpha)$ as before, it follows from Lemma 45 (Section 13) that V *is* a null set on U_* with respect to the measure du_*. Similarly $\bigcup_{\alpha \varepsilon \Sigma^+} \mathfrak{A}^*_\alpha$ is a null set on $\mathfrak{A}^*$ with respect to the measure da. Since $\psi(\bar{K}_0 \times \mathfrak{A}^{*\prime}) = {}^cV$, our assertion follows from Corollary 1 above.

Put $J^* = K_0 \cap \mathfrak{A}^*$. Since $\mathfrak{k}_0 \cap (-1)^{\frac{1}{2}} \mathfrak{a}_{\mathfrak{p}_0} = 0$, J^* is both discrete and compact and so it is finite. It is obvious that $(J^*)^s = J^*$ $(s \varepsilon W)$.

LEMMA 43. *Let B^*_0 be any connected component of $\mathfrak{A}^{*\prime}$. Then*

$$\mathfrak{A}^{*\prime} = \bigcup_{s \varepsilon W} (B^*_0)^s J^*.$$

Hence $\mathfrak{A}^{*\prime}$ has only a finite number of connected components.

Let $\bar{B}^*_0$ denote the closure of B^*_0 in $\mathfrak{A}^*$. Since B^*_0 is closed in $\mathfrak{A}^{*\prime}$, $B^*_0 = \bar{B}^*_0 \cap \mathfrak{A}^{*\prime}$. Now define V and ${}^cV$ as above. Then

$$\psi(\bar{K}_0 \times B^*_0) = {}^cV \cap \psi(\bar{K}_0 \times \bar{B}^*_0)$$

is both open and closed in ${}^cV$ and therefore, ${}^cV$ being connected,

$${}^cV = \psi(\bar{K}_0 \times B^*_0) \text{ and } U_* = \psi(\bar{K}_0 \times \bar{B}^*_0).$$

Now suppose $a \varepsilon \mathfrak{A}^{*\prime}$. Then we can find elements $b \varepsilon \bar{B}^*_0$ and $k \varepsilon K_0$ such that $a_* = (kb)_*$. Hence $(k^{-1}a)_* = b_*$ and therefore $k \varepsilon M'$ from Lemma 42. This means that $b_* = (a^s)_*$ for some $s \varepsilon W$ and $b^{-1}a^s \varepsilon K_0 \cap \mathfrak{A}^* = J^*$. Conversely if $z \varepsilon J^*$, $z^{-1} = \theta(z)$ and so $z^2 = 1$. Hence $|D(az)| = |D(a)|$ for any $a \varepsilon \mathfrak{A}^*$. This shows that $\bigcup_{s \varepsilon W} (B^*_0)^s J^*$ is contained in $\mathfrak{A}^{*\prime}$ and so the lemma follows.

COROLLARY. *Let w^* be the number of connected components of $\mathfrak{A}^{*\prime}$. Then if f is any continuous function on U,*

$$\int_U f(u)\, du \int_{\mathfrak{A}^*} |D(a)|^{\frac{1}{2}} da = w^* \int_{B_0^*} |D(a)|^{\frac{1}{2}} da \int_{K_0 \times K_0} f(kak')\, dk dk'.$$

Let f^* denote the function on U_* given by

$$f^*(u_*) = \int_{K_0} f(uk)\, dk \qquad (u \varepsilon U).$$

Then

$$\int_{U_*} f^*(u_*)\, du_* = \int_U f(u)\, du$$

and

$$\int_{K_0} f(kak')\, dk' = f^*((ka)_*).$$

Therefore

$$\int_{K_0 \times K_0} f(kak')\,dkdk' = \int_{K_0} f^*((ka_*)\,dk$$

and so it follows from Corollary 2 to Lemma 42 that

$$\int_U f(u)\,du = c\int_{\mathfrak{A}^*} |D(a)|^{\frac{1}{2}}\,da \int_{K_0 \times K_0} f(kak')\,dkdk'.$$

Now we can replace the integral on $\mathfrak{A}^*$ on the right by the corresponding integral on $\mathfrak{A}^{*\prime}$. We know that any component B^* of $\mathfrak{A}^{*\prime}$ is of the form $(B^*{}_0)^s z$ $(s \varepsilon W, z \varepsilon J^*)$. Therefore if m is any element of M' in the coset s, it is clear that

$$\begin{aligned}\int_{B^*} |D(a)|^{\frac{1}{2}}\,da \int_{K_0 \times K_0} f(kak')\,dkdk' &= \int_{B_0^*} |D(a^s z)|^{\frac{1}{2}}\,da \int_{K_0 \times K_0} f(kmam^{-1}zk')\,dkdk' \\ &= \int_{B_0^*} |D(a)|^{\frac{1}{2}}\,da \int_{K_0 \times K_0} f(kak')\,dkdk'\end{aligned}$$

since $|D(a^s z)| = |D(a))$. This shows that

$$\int_{\mathfrak{A}^*} |D(a)|^{\frac{1}{2}}\,da \int_{K_0 \times K_0} f(kak')\,dkdk' = w^* \int_{B_0^*} |D(a)|^{\frac{1}{2}}\,da \int_{K_0 \times K_0} f(kak')\,dkdk'.$$

In particular if we take the function $f = 1$, we get

$$1 = \int_U du = c\int_{\mathfrak{A}^*} |D(a)|^{\frac{1}{2}}\,da$$

and this gives our result.

Now if we normalize the Haar measure da on $\mathfrak{A}^*$ in accordance with Lemma 22, it follows from Lemma 39 that the numbers w^* and $\int_{\mathfrak{A}^*} |D(a)|^{\frac{1}{2}}\,da$ are independent of the choice of $\mathfrak{a}_{\mathfrak{p}_0}$. This completes the proof of Lemma 22.

13. Appendix. We shall now give a proof of Lemma 32. Let m and n be the dimensions of M and N respectively and let $(t_1, \cdots, t_m)$ be a coordinate system on M which is valid on some open neighborhood V of a point $x_0 \varepsilon M$. We assume $t_i(x_0) = 0$ $i = 1, \cdots, m$. For any positive integer p let R^p denote the Cartesian product of R with itself p times. It is obvious that we can chose a compact neighborhood W of x_0 and a positive number a such that $W \subset V$ and the mapping $x \to (t_1(x), \cdots, t_m(x))$ $(x \varepsilon W)$ maps W topologically on the cube $|t_i| \leqq a$ in R^m. Any set W defined in this way

is called a cubic set in M (with respect to the coordinate system $(t_1,\cdots,t_m)$). Cubic sets in N are defined similarly.

Since M satisfies the countability axioms (see Chevalley [3]) we can find a countable family V_i $(i=1,2,\cdots)$ of cubic sets in M such that the following conditions hold: (1) $M=\bigcup_i V_i$ and (2) for each i, $f(V_i)$ is contained in some cubic set in N. Therefore, in view of the sum theorem of dimension theory (Hurewicz and Wallman [6, p. 30]) it is enough to show that $\partial\mathrm{im}\, f(V_i)\leqq m$ for each i. This however is an immediate consequence of the lemma below.

We regard R^p as an additive group in the usual way and if

$$a=(a_1,\cdots,a_p)\,\varepsilon\, R^p,\text{ we put } |a|=(\sum_{1\leqq i\leqq p}a_i^2)^{\frac{1}{2}}.$$

Let I be the unit interval $0\leqq t\leqq 1$ in R and let I^m be the Cartesian product of I with itself m times. Then $I^m\subset R^m$.

LEMMA 44. *Let f be a mapping of I^m into R^n such that*

$$|f(b)-f(a)|\leqq c\,|b-a| \qquad (a,b\,\varepsilon\, I^m)$$

where c is a fixed number. Then $\partial\mathrm{im}\, f(I^m)\leqq m$.

It is enough to prove that the Hausdorff $(m+1)$-measure of $f(I^m)$ is zero (see Hurewicz and Wallman [6, p. 104]). But this follows at once from our hypothesis on f and the fact that the Hausdorff $(m+1)$-measure of I^m is zero (Hurewicz and Wallman [6, p. 103]).

Let μ be a Borel measure on N. We say that μ is locally euclidean if the following condition holds. For each x_0 in N we can find a cubic set W in N with respect to a coordinate system $(t_1,\cdots,t_n)$ such that (1) x_0 lies in the interior of W and (2) μ is completely continuous (on W) with respect to the Euclidean measure $dt=dt_1\cdots dt_n$ on W.

LEMMA 45. *Let N be a manifold satisfying the countability axioms and let μ be a locally euclidean measure on N. Then if Q is any subset of N with $\partial\mathrm{im}\, Q<\partial\mathrm{im}\, N$, Q is a null set with respect to μ.*

We can cover N by a countable family of cubic sets V_i $i=1,2,\cdots$. For any fixed i, choose a coordinate system $(t_1,\cdots,t_n)$ valid on some open neighborhood of V_i such that V_i is a cubic set with respect to this system. Obviously it would be enough to prove that $Q\cap V_i$ is a null set with respect to the measure $dt=dt_1\cdots dt_n$. So we have to prove the following lemma.

LEMMA 46. *Let Q be a subset of I^n such that* $\partial\mathrm{im}\, Q < n$. *Then Q is a null set with respect to the Euclidean measure on I^n.*

Since $\partial\mathrm{im}\, Q < n$, the Hausdorff n-measure of Q is zero. Our assertion therefore follows immediately from the definition of the Hausdorff measure.

THE INSTITUTE FOR ADVANCED STUDY,
PRINCETON, N. J.
COLUMBIA UNIVERSITY,
NEW YORK, N. Y.

REFERENCES.

[1] V. Bargmann, *Ann. of Math.*, vol. 48 (1947), pp. 568-640.

[2] E. Cartan, (a) *Ann. École Norm. Sup.*, vol. 44 (1927), pp. 345-467.
(b) *J. Math. Pures Appl.*, vol. 8 (1929), pp. 1-33.
(c) *Abh. Math. Seminar Hamburg*, vol. 11 (1925), pp. 116-162.

[3] C. Chevalley, *Theory of Lie groups*, Princeton University Press, 1946.

[4] R. Godement, *C. R. Acad. Sci. Paris*, vol. 225 (1947), (a) pp. 521-523, (b) pp. 657-659.

[5] Harish-Chandra. (a) *Trans. Amer. Math. Soc.*, vol. 70 (1951), pp. 28-96.
(b) vol. 75 (1953), pp. 185-243.
(c) vol. 76 (1953), pp. 234-253.
(d) *Proc. Nat. Acad. Sci.*, vol. 41 (1955), pp. 314-317.
(e) *Amer. Jour. Math.*, vol. 77 (1955), pp. 743-777.
(f) *Amer. Jour. Math.*, vol. 78 (1956), pp. 1-41.

[6] W. Hurewicz and H. Wallman, *Dimension theory*, Princeton University Press, 1948.

[7] J. L. Koszul, *Bull. Soc. Math.*, vol. 78 (1950), pp. 65-127.

[8] F. I. Mautner, *Ann. of Math.*, vol. 52 (1950), pp. 528-556.

[9] G. D. Mostow, *Bull. Amer. Math. Soc.*, vol. 55 (1949), pp. 969-980.

[10] I. E. Segal, (a) *Ann. of Math.*, vol. 52 (1950), pp. 272-292.
(b) *Proc. Amer. Math. Soc.*, vol. 3 (1952), pp. 13-15.

[11] H. Weyl, (a) *Math. Zeit.*, vol. 24 (1925), pp. 328-395.
(b) The structure and representations of continuous groups, Princeton, The Institute for Advanced Study, 1935.

LIE AND JORDAN SYSTEMS IN SIMPLE RINGS WITH INVOLUTION.*

By I. N. HERSTEIN.

In previous papers, ([3], [4], [5]), the author has recently considered the question of Lie and Jordan simplicity of simple, associative rings and of certain subsets thereof. Baxter [1] completed certain cases, namely characteristic 2 and 3. These considerations were motivated by the classical results (in the theory of simple Lie algebras) which yielded that total matrix algebras over fields of charateristic 0 gave rise, in a natural way, to simple Lie algebras.

However, in the total matrix algebras over fields, adjoints (involutions) can be defined in a variety of ways, and it was also known classically that the skew elements under the adjoint also led to simple Lie algebras; more recently, it has been demonstrated that the self-adjoint elements yield simple Jordan algebras. With these examples as motivation, we proceed, in this paper, to study the general situation, namely the Lie and Jordan structure of the skew and self-adjoint elements of arbitrary simple rings possessing involutions.

In the course of this study several other results, interesting in their own rights, fall our way. For instance we prove (Theorem 9) that the subring generated by the self-adjoint elements is A except in the 4-dimensional case; this generalizes a result of Dieudonné [2]. We also show that under the same conditions, as above, the skew elements generate the full ring (Theorem 15); this is considerably more difficult than the case of the self-adjoint elements. There are also some special results about representations in various forms of self-adjoint elements and also concerning the taking of commutators twice of the skew elements.

1. Preliminaries. Let A be a simple ring of characteristic different from 2. We say a mapping, $*$, from A to A is an adjoint or involution on A if

* Received November 14, 1955.

(1) $a^{**}=a$, (2) $(a+b)^*=a^*+b^*$, (3) $(ab)^*=b^*a^*$,

for all a, b in A.

As usual, we define S, the set of self-adjoint elements of A, by $S=\{x\varepsilon A \mid x^*=x\}$. Clearly S is an additive subgroup of A. Moreover, if $a,b\varepsilon S$ then $ab+ba\varepsilon S$. Thus S is a Jordan subring of A under the Jordan product which is defined in A by $x\circ y=xy+yx$ for $x,y\varepsilon A$. An additive subgroup, U, of S is said to be a *Jordan ideal* of S if whenever $u\varepsilon U$, $x\varepsilon S$ then $ux+xu\varepsilon U$.

We further define K, the set of skew elements of A, by $K=\{x\varepsilon A \mid x^* =-x\}$. As is readily verified, K is an additive subgroup of A, and if $x,y\varepsilon K$ then $xy-yx\varepsilon K$. Thus K is a Lie subring of A under the Lie product $[x,y]$ defined in A by $[x,y]=xy-yx$ for all $x,y\varepsilon A$. We say that an additive subgroup, U, of K is a *Lie ideal* of K if $x\varepsilon K$, $u\varepsilon U$ implies that $xu-ux\varepsilon U$.

Certain facts about adjoints are trivial and we shall make use of these throughout the body of this paper without proof, explanation or reference; such facts include, for instance, $x^*-x\varepsilon K$, $x^*+x\varepsilon S$, $s\varepsilon S$, $k\varepsilon K$ then $sk+ks\varepsilon K$, $sk-ks\varepsilon S$, etc.

2. Jordan simplicity of S. The purpose of this section is to show that the results from matrix theory extend to general simple rings, that is, that S is a simple Jordan ring. Throughout this section U will denote a non-zero Jordan ideal of S. Our purpose is to show that $U=S$.

LEMMA 1. *If $x,y\varepsilon S$ and $u\varepsilon U$ then $(xu-ux)y-y(xu-ux)\varepsilon U$.*

Proof. Since U is a Jordan ideal of S and since $xy+yx\varepsilon S$, it follows that $(xy+yx)u+u(xy+yx)\varepsilon U$. However,

$$(xy+yx)u+u(xy+yx)$$
$$=\{x(yu+uy)+(yu+uy)x\}+\{(ux+xu)y-y(ux-xu)\}.$$

Now $yu+uy\varepsilon U$ since U is a Jordan ideal of S; thus, by the same token, $x(uy+yu)+(uy+yu)x\varepsilon U$. That is, the first $\{\ \}$ on the right-hand side of the relation above is in U. Since the left-hand side of this identity is also in U, we are left with the fact that the second $\{\ \}$ on the right-hand side must be in U. That is, $(ux-xu)y-y(ux-xu)$ is in U, which is the required lemma.

LEMMA 2. *If $u\varepsilon U$, $b\varepsilon K$ then $bu^2-u^2b\varepsilon U$.*

For, $ub - bu \in S$, and so $(ub - bu)u + u(ub - bu) \in U$; since

$$(ub - bu)u + u(ub - bu) = u^2b - bu^2$$

this completes the proof of the lemma.

LEMMA 3. *If $u \in U$, $x \in S$ then $xux \in U$.*

Proof. $2xux = \{x(xu + ux) + (xu + ux)x\} - \{x^2u + ux^2\}$. Since U is a Jordan ideal of S, and since $x^2 \in S$ along with x, both $\{\ \}$ on the right-hand side are in U, and so $2xux \in U$. Thus $2(2xux) \in U$; that is $(2x)u(2x) \in U$ for all $x \in S$. Since the characteristic of A is not 2, it is easily seen that $2S = S$; hence $(2x)u(2x) \in U$ for all $x \in S$ implies $xux \in U$ for all $x \in S$.

COROLLARY. *If $x, y \in S$ and $u \in U$ then $xuy + yux \in U$.*

For, linearizing the result of the lemma by replacing x by $x + y$ the corollary follows immediately from the lemma.

Our aim is to show that $au^4b + b^*u^4a^*$ is in U for all $a, b \in A$. We have, in the corollary above, disposed of the case $a^* = a$, $b^* = b$. We proceed in the next few lemmas to consider the other special possibilities, $a^* = -a$, $b^* = +b$, $a^* = -a$, $b^* = -b$, and then to combine them for the result at the opening of this paragraph. We now show

LEMMA 4. *If $u \in U$, $b \in S$, $a \in K$ then $bu^2a - au^2b \in U$.*

Proof. By Lemma 2, since $a \in K$, $u^2a - au^2 \in U$. Then, U being a Jordan ideal of S, $b(u^2a - au^2) + (u^2a - au^2)b \in U$. Now

(1) $\qquad b(u^2a - au^2) + (u^2a - au^2)b = bu^2a - au^2b - bau^2 + u^2ab.$

Consider

(2) $\qquad a(u^2b - bu^2) + (u^2b - bu^2)a = au^2b - bu^2a - abu^2 + u^2ba.$

Adding (1) and (2), the right-hand sides add up to $u^2(ab + ba) - (ab + ba)u^2$; since $ab + ba \in K$, by Lemma 2, $u^2(ab + ba) - (ab + ba)u^2$ is in U. Thus the sum of the right-hand sides of (1) and (2) is in U; consequently the sum of the left-hand sides is also in U. Since the left-hand side of (1) is already in U, we obtain that the left-hand side of (2) must also be in U. Thus, if we now subtract (1) from (2) we observe that the elements we get on each side of the resulting equation are in U. In particular, the difference of the right-hand sides is an element of U; whence

(3) $\qquad 2(au^2b - bu^2a) + (ba - ab)u^2 + u^2(ba - ab) \in U.$

Now $2u^2 = uu + uu \varepsilon U$, so $2u^2s + s2u^2 = u^2(2s) + (2s)u^2 \varepsilon U$ for all $s \varepsilon S$, and since $2S = S$, we have $u^2s + su^2 \varepsilon U$. In particular, in (3), since $ab - ba \varepsilon S$, $u^2(ab - ba) + (ab - ba)u^2 \varepsilon U$; from which (3) reduces to $2(au^2b - bu^2a) = au^2(2b) - (2b)u^2a \varepsilon U$. Since $2S = S$ this gives rise to $au^2b - bu^2a \varepsilon U$ for all $a \varepsilon K$, $b \varepsilon S$, $u \varepsilon U$, which is Lemma 4.

LEMMA 5. *If $a \varepsilon K$, $u \varepsilon U$ then $au^4a \varepsilon U$.*

Proof. By Lemma 2, $au^2 - u^2a \varepsilon U$. Thus $2(au^2 - u^2a)^2 \varepsilon U$, and so $4(au^2 - u^2a)^2 = ((2a)u^2 - u^2(2a))^2 \varepsilon U$. Since $2K = K$, (as is easily seen) we obtain that $(au^2 - u^2a)^2 \varepsilon U$ for all $a \varepsilon K$, $u \varepsilon U$. But

$$(au^2 - u^2a)^2 = \{(au^2a)u^2 + u^2(au^2a)\} - u^2a^2u^2 - au^4a.$$

Since $au^2a \varepsilon S$, as we saw in the proof of Lemma 4, $(au^2a)u^2 + u^2(au^2a) \varepsilon U$. Also

$$2u^2a^2u^2 = \{(u^2a^2 + a^2u^2)u^2 + u^2(u^2a^2 + a^2u^2)\} - \{u^4a^2 + a^2u^4\},$$

and since $a^2 \varepsilon S$, both $\{\ \}$ are in U, so $2u^2a^2u^2 \varepsilon U$, from which, as before, $u^2a^2u^2 \varepsilon U$. Thus we obtain that $au^4a \varepsilon U$.

Linearizing the lemma we have

COROLLARY. *If $a, b \varepsilon K$ then for $u \varepsilon U$, $au^4b + bu^4a \varepsilon U$.*

We are now able to prove

THEOREM 6. *If $u \varepsilon U, r, t \varepsilon A$ then $ru^4t + t^*u^4r^* \varepsilon U$.*

Proof. Clearly $r = r_0 + r_1$ where $r_0 \varepsilon S$, $r_1 \varepsilon K$ $(r_0 = \frac{1}{2}(r^* + r)$, $r_1 = \frac{1}{2}(r - r^*))$, and $t = t_0 + t_1$ where $t_0 \varepsilon S$, $t_1 \varepsilon K$. Thus

$$\begin{aligned} ru^4t + t^*u^4r^* &= (r_0 + r_1)u^4(t_0 + t_1) + (r_0 - r_1)u^4(t_0 - t_1) = (r_0u^4t_0 + t_0u^4r_0) \\ &+ (r_0u^4t_1 - t_1u^4r_0) + (r_1u^4t_0 - t_0u^4r_1) + (r_1u^4t_1 + t_1u^4r_1). \end{aligned}$$

By the corollary to Lemma 5, since $r_1, t_1 \varepsilon K$, $(r_1u^4t_1 + t_1u^4r_1) \varepsilon U$. Since $2u^2 \varepsilon U$, $4u^4 \varepsilon U$, and since $2S = S$, by Lemma 3, $(r_0u^4t_0 + t_0u^4r_0) \varepsilon U$. Since $2u^2 \varepsilon U$, and since $2S = S$, by Lemma 4, $(r_0u^4t_1 - t_1u^4r_0) \varepsilon U$ and likewise $(r_1u^4t_0 - t_0u^4r_1) \varepsilon U$. Thus each component piece in the expression for $ru^4t + t^*u^4r^*$ is in U, so $ru^4t + t^*u^4r^* \varepsilon U$ for all $r, t \varepsilon A$.

THEOREM 7. *If $U \neq S$ then $u \varepsilon U$ implies that $u^4 = 0$.*

Proof. Suppose that $u^4 \neq 0$ for some $u \varepsilon U$. Since A is a simple ring, $Au^4A = A$. Thus if $y \varepsilon A$, $y = \sum r_iu^4t_i$ where $r_i, t_i \varepsilon A$. Hence $y^* = \sum t_i^*u^4r_i^*$. But then $y + y^* = \sum (r_iu^4t_i + t_i^*u^4r_i^*)$ and so in U since each $r_iu^4t_i + t_i^*u^4r_i^*$

is in U by Theorem 6. That is $y+y^* \varepsilon U$ for every $y \varepsilon A$. Since every element of S is so representable we obtain $U=S$, a contradiction. In this way we are forced to $u^4=0$ for all $u \varepsilon U$.

We are now able to prove the principal theorem of this section.

THEOREM 8. *If A is a simple ring of characteristic $\neq 2$ then S is a simple Jordan ring.*

Proof. Suppose U is a Jordan ideal of S and that $U \neq S$. If $r \varepsilon A$, $u \varepsilon U$ and $r=r_0+r_1$, $r_0 \varepsilon S$, $r_1 \varepsilon K$ then

$$u^2r+r^*u^2=u^2(r_0+r_1)+(r_0-r_1)u^2=(u^2r_0+r_0u^2)+(u^2r_1-r_1u^2),$$

and so is in U by Lemma 2. Consequently by the previous theorem, $(u^2r+r^*u^2)^4=0$. Since $u^4=0$, multiplying $(u^2r+r^*u^2)^4=0$ by u^2 from the right and r from the left we have $r(u^2r+r^*u^2)^4u^2=0$; simplifying this we obtain $(ru^2)^5=0$. But then Au^2 is a left-ideal all of whose elements are nilpotent of index 5; by a theorem of Levitzki [9], Au^2 must be locally nilpotent. If $Au^2 \neq 0$, by another theorem of Levitzki [10], A must possess a non-zero locally nilpotent two-sided ideal. The simplicity of A then forces A to be locally nilpotent; this is impossible in a simple ring [3]. Thus $Au^2=(0)$ and because A is simple, $u^2=0$ results. That is, $u^2=0$ for every $u \varepsilon U$. Therefore for every $u, v \varepsilon U$, $uv+vu=(u+v)^2-u^2-v^2=0$. If $x \varepsilon S$, $v=xu+ux \varepsilon U$, and so $u(xu+ux)+(xu+ux)u=0$. This reduces to $2uxu=0$ since $u^2=0$. A has characteristic $\neq 2$, thus $uxu=0$ for all $x \varepsilon S$. If $a \varepsilon K$, $aua \varepsilon S$, so by the above $uauau=u(aua)u=0$. If r is any element of A, $r=r_0+r_1$, where $r_0 \varepsilon S$, $r_1 \varepsilon K$; whence

$$ururu=u(r_0+r_1)u(r_0+r_1)u=ur_1ur_1u=0.$$

In this way Au is a left-ideal in which every element is nilpotent of index 3. Using the argument as above we are led to $Au=(0)$, and so $u=0$. That is $U=(0)$. Thus the only proper Jordan ideal of S is (0), consequently S is a simple Jordan ring.

The following theorem is a generalization to the case of an arbitrary simple ring of a result Dieudonné proved for division rings [2]. We shall need it for use later in this paper when we study the Lie ideal structure of K.

We first define: If B is a subset of A then $\bar{B}$ is the subring of A generated by B.

THEOREM 9. *If A is a simple ring of characteristic $\neq 2$ and if Z, the center of A, is (0) or if A is more than 4-dimensional over Z, then $\bar{S}=A$.*

Proof. By its very definition $\bar{S}$ is a subring of A. We claim that $\bar{S}$ is, in addition, a Lie ideal of A. For, if $a \varepsilon \bar{S}$, $s \varepsilon S$, clearly $as - sa \varepsilon \bar{S}$. On the other hand, if $b, a \varepsilon S$, $k \varepsilon S$, then $abk - kab = a(bk - kb) + (ak - ka)b$ and so is in $\bar{S}$ since $[K, S] \subset S$; continuing in this way, we have $ak - ka \varepsilon \bar{S}$ for all $a \varepsilon \bar{S}$, $k \varepsilon K$. Since every $r \varepsilon A$ can be written as $r = s + k$ with $s \varepsilon S$ and $k \varepsilon K$, we obtain that $ar - ra \varepsilon \bar{S}$ for all $a \varepsilon \bar{S}$ and all $r \varepsilon A$. Thus $\bar{S}$ is a Lie ideal of A. By Theorem 3 of [3] either $\bar{S} = A$ or $\bar{S} \subset Z$, the center of A. Suppose $\bar{S} \neq A$; then the other possibility, $\bar{S} \subset Z$, must prevail. In particular $S \subset Z$. Given $r \varepsilon A$, $r = k + \lambda$ where $\lambda \varepsilon S \subset Z$, $k \varepsilon K$. Thus $(r - \lambda)^2 = k^2 \varepsilon Z$ since $k^2 \varepsilon S$. Therefore every element of A satisfies a quadratic equation over Z. In this discussion $Z \neq (0)$ (otherwise $\bar{S} \subset Z$ is nonsensical), so A is a primitive ring, in which every element satisfies an equation of degree 2 over the center. By the work of Jacobson [6] this implies that A is at most 4-dimensional over Z. The proof of Theorem 9 is thereby completed.

3. The Lie ideals of *K*. We now turn our attention to the problem of finding the possible Lie ideals of K. Let U be a Lie ideal of K; our aim is to show that either $U \subset Z$ or $U \supset [K, K]$. Our starting point in these considerations is

LEMMA 10. *If $u \varepsilon U$, $x \varepsilon S$ then $u^2x - xu^2 \varepsilon U$.*

Proof. Since $u \varepsilon U \subset K$, $x \varepsilon S$, it follows that $xu + ux \varepsilon K$. Consequently $u^2x - xu^2 = (xu + ux)u - u(xu + ux) \varepsilon U$ since U is a Lie ideal of K.

In the case of Jordan ideals we had that u in the ideal implies that $2u^2$ is also in that ideal. In the Lie case we have no such closure under squaring, but the next lemma provides a substitute, showing that although u^3 need not be in U although $u \varepsilon U$, u^3 does have a close relation to U.

LEMMA 11. *If $u \varepsilon U$ and $x \varepsilon K$ then $u^3x - xu^3 \varepsilon U$; that is $[u^3, K] \subset U$.*

Proof. Since $xu + ux \varepsilon S$, by Lemma 10, $(xu + ux)u^2 - u^2(xu + ux) \varepsilon U$. In other words, $xu^3 - u^3x + uxu^2 - u^2xu \varepsilon U$. However, $uxu \varepsilon K$, from which it follows that $uxu^2 - u^2xu = uxu \cdot u - u \cdot uxu \varepsilon U$. By the above we are left with $xu^3 - u^3x$ completing the proof.

For U a Lie ideal of K, we define $T(U)$ by $T(U) = \{x \varepsilon K \mid [x, K] \subset U\}$. We note some properties of $T(U)$ in

LEMMA 12. *$T(U)$ is a Lie ideal of K; moreover $U \subset T(U)$.*

Proof. That $U \subset T(U)$ is, of course, nothing more than the definition of a Lie ideal of K. It is also clear that $T(U)$ is an additive subgroup of U.

To prove the lemma there remains but to show that $[T(U),K]\subset T(U)$. Since $[T(U),K]\subset U$ by definition, and since $U\subset T(U)$, this fact follows easily also.

Although $u\in U$ need not imply that $u^3\in U$ we note

LEMMA 13. *$t\in T(U)$ implies that $t^3\in T(U)$.*

Proof. Suppose that $t\in T(U)$, $x\in K$. We must show that $t^3x-xt^3\in U$. Now if $s\in S$, $t^2s-st^2=t(ts+st)-(ts-st)t$ is in U since $ts+st\in K$. In particular, for $s=tx+xt$,

$$U\ni t^2(tx+xt)-(tx+xt)t^2=t^3x-xt^3+t\cdot txt-txt\cdot t.$$

Since $txt\in K$, $txt\cdot t-t\cdot txt\in U$; whence $t^3x-xt^3\in U$ and $t^3\in T(U)$.

Although we do not need the concept in this paper, we feel that it will prove useful in further considerations on the Lie structure of rings with involutions, so we define:

U is a *strong Lie ideal* of K if $u\in U$ implies that $u^3\in U$.

We note the following corollary to Lemma 13:

COROLLARY. *If $U\not\supset[K,K]$ is a Lie ideal of K then U can be imbedded in a proper, strong Lie ideal of K.*

For if $U\not\supset[K,K]$, $T(U)\neq K$. Since, by Lemma 13, $T(U)$ is a strong Lie ideal of K and $U\subset T(U)$ the corollary follows.

Let $K\circ K$ be the additive subgroup generated by all $xy+yx$ where x and y range over K. Clearly $K\circ K\subset S$.

LEMMA 14. *$S=[K,S]+K\circ K$.*

Proof. As noted above, $K\circ K\subset S$. It is equally trivial that $[K,S]\subset S$. We claim that $[K,S]+K\circ K$ is a Jordan ideal of S. It is obviously an additive subgroup of S. Suppose now that $a,b\in K$ and $s\in S$. Thus

$$(ab+ba)s+s(ab+ba)$$
$$=\{a(bs-sb)-(bs-sb)a\}+\{(as+sa)b+b(as+sa)\}.$$

The first $\{\ \}$ is in $[K,S]$ since $bs-sb\in S$; since $as+sa\in K$, the second $\{\ \}$ is in $K\circ K$. Consequently $(K\circ K)\circ S\subset[K,S]+K\circ K$.

On the other hand, if $a\in K$, $s,t\in S$ then

$$(as-sa)t+t(as-sa)$$
$$=\{a(st-ts)+(st-ts)a\}+\{(at+ta)s-s(at+ta)\}.$$

Since $st - ts \in K$, the first $\{\ \}$ is in $K \circ K$; since $at + ta \in K$, the second $\{\ \}$ is in $[K, S]$. Thus $[K, S] \circ S \subset [K, S] + K \circ K$.

In other words, $[K, S] + K \circ K$ is a Jordan ideal of S. By Theorem 8 it follows that $S = [K, S] + K \circ K$, (the desired result) or $[K, S] + K \circ K = (0)$. We wish to rule out the second possibility. If $[K, S] + K \circ K = (0)$, then in particular, $k \in K$ implies that $k^2 = 0$. Also, $k \in K$ and $s \in S$ implies that $ks = sk$. Thus $0 = k^2 s = ksk$; that is $kSk = (0)$. If, on the other hand, $a \in K$, then $ka + ak = 0$ since it is in $K \circ K = (0)$; multiplying this on the right by k we obtain $kak = 0$; hence $kKk = (0)$. Since $A = K + S$, $kAk = k(K + S)k = kKk + kSk = (0)$. kA is, in this way, a nilpotent right ideal, which is impossible in a simple ring unless $k = 0$. Since $K \neq (0)$ (otherwise A is a commutative field), $[K, S] + K \circ K = (0)$ is not a possibility, and so it is equal to the only other possibility, namely S.

We recall that $\bar{K}$ is the subring of A generated by K. Although the next theorem is of some independent interest, it is the essential key to all the results that follow.

THEOREM 15. *If A has a trivial center or if A is more than 4-dimensional over its center then $\bar{K} = A$.*

Proof. By definition, $K \subset \bar{K}$. We now consider those cases for which $S \subset \bar{K}$ will be provable.

Let $a, b \in K$, and let $s \in S$. Thus $sa + as \in K$, and so $(sa + as)b \in \bar{K}$. That is, $sab + asb \in \bar{K}$. Similarly $sba + bsa \in \bar{K}$. Subtracting these two we obtain that $s(ab - ba) + asb - bsa \in \bar{K}$. However, $asb - bsa = asb - b^* s^* a^* = asb - (asb)^*$ and thus is in K, and so in $\bar{K}$. Consequently we have demonstrated that $s(ab - ba) \in \bar{K}$. Rephrasing this, $S(ab - ba) \subset \bar{K}$. However, since $ab - ba \in K$, $K(ab - ba) \subset \bar{K}$, and since $A = K + S$, we obtain that $A(ab - ba) \subset \bar{K}$ for all $a, b \in K$. Similarly, if $c, d \in K$, $(cd - dc)A \subset \bar{K}$. Since $\bar{K}$ is a subring of A, and since both $A(ab - ba)$ and $(cd - dc)A$ are contained in $\bar{K}$, $A(ab - ba)(cd - dc)A \subset \bar{K}$. A is a simple ring and $A(ab - ba)(cd - dc)A$ is a two-sided ideal of A, so either $A(ab - ba)(cd - dc)A = A$, in which case $A \subset \bar{K}$, the desired result, or $A(ab - ba)(cd - dc)A = (0)$ for all $a, b, c, d \in K$. We consider when the second possibility can hold. In that case, by the simplicity of A, $(ab - ba)(cd - dc) = 0$ for all $a, b, c, d \in K$. Since $K(cd - dc)A \subset \bar{K}$, the same argument as used above leads to $(ab - ba)\bar{K}(cd - dc) = (0)$ for all $a, b, c, d \in K$. Suppose now that $s \in S$, $c, d \in K$. Then

$$s(cd + dc) - (cd + dc)s$$
$$= \{(sc + cs)d - d(sc + cs)\} + \{(ds + sd)c - c(ds + sd)\}.$$

Since $sc+cs\,\varepsilon\,K$, the first { } is in $[K,K]$; similarly the second { } is in $[K,K]$. Thus the left-hand side is in $[K,K]$. But then, since we know from the argument above that $(ab-ba)[K,K]=(0)$ for $a,b\,\varepsilon\,K$, we have that $(ab-ba)\{s(cd+dc)-(cd+dc)s\}=(0)$ for all $a,b,c,d\,\varepsilon\,K$ and $s\,\varepsilon\,S$. If we further suppose that $d=ef-fe$ where $e,f\,\varepsilon\,K$ then $(ab-ba)dc=0$ and $(ab-ba)cd=0$ since it is contained in $(ab-ba)\bar{K}d=(0)$; and so $(ab-ba)(cd+dc)=0$. Thus $(ab-ba)\{(cd+dc)s-s(cd+dc)\}=0$ reduces to $(ab-ba)s(cd+dc)=0$ for all $s\,\varepsilon\,S$, when $d=ef-fe$. That is, $(ab-ba)S(c(ef-fe)+(ef-fe)c)=(0)$ for all $a,b,c,e,f\,\varepsilon\,K$. Since, in addition, $(ab-ba)K(c(ef-fe)+(ef-fe)c)=(0)$ (since it is contained in $(ab-ba)\bar{K}(ef-fe)=(0)$) and since $A=K+S$, we obtain $(ab-ba)A(c(ef-fe)+(ef-fe)c)=(0)$ for all $a,b,c,e,f\,\varepsilon\,K$. We wish to show that $ab=ba$ for all $a,b\,\varepsilon\,K$. If not, that is, if $ab-ba\neq 0$ for some $a,b\,\varepsilon\,K$, then by the simplicity of A, $c(ef-fe)+(ef-fe)c=0$ for all c,e,f in K. If $s\,\varepsilon\,S$ then $c=(ef-fe)s+s(ef-fe)$ is in K and so $c(ef-fe)+(ef-fe)c=0$ implies, since $(ef-fe)^2=0$, that $2(ef-fe)s(ef-fe)=0$; that is, $(ef-fe)S(ef-fe)=(0)$. Since, from before, $(ef-fe)K(ef-fe)=(0)$ we are led to $(ef-fe)A(ef-fe)=(0)$, which forces $ef-fe=0$ by the simplicity of A. So, $\bar{K}\neq A$ has resulted in $ab=ba$ for all $a,b\,\varepsilon\,K$. In particular, if $a\,\varepsilon\,K$, a^2 commutes with all elements of K. Now if $a\,\varepsilon\,K$, $s\,\varepsilon\,S$, then $as+sa\,\varepsilon\,K$, and so $a(as+sa)=(as+sa)a$; in simplifying this says that $a^2s=sa^2$ for all $s\,\varepsilon\,S$. Since a^2 commutes with all elements of K and of S, and since $K+S=A$, a^2 commutes with all elements of A. Consequently $a^2\,\varepsilon\,Z$, the center of A, for all $a\,\varepsilon\,K$. Linearizing this we obtain that $ab+ba$ is in Z for all $a,b\,\varepsilon\,K$, and so $K\circ K\subset Z$. Since $ab=ba$, we have that $2ab=ab+ba\,\varepsilon\,Z$. Thus $a,b\,\varepsilon\,K$ implies that $ab\,\varepsilon\,Z$. Since $Z=(0)$ or Z is a field, $ab=0$ or a has an inverse in A for all $a,b\,\varepsilon\,K$. If a has an inverse for some $a\,\varepsilon\,K$ then so does b for every $b\neq 0\,\varepsilon\,K$, for $ab\neq 0\,\varepsilon\,Z$ and so has an inverse, whence b has an inverse in A. We rule out the possibility that no $a\,\varepsilon\,K$ has an inverse in A. For then, since $a^2\,\varepsilon\,Z$, $a^2=0$ results; also, since $ab\,\varepsilon\,Z$, $ab=0$ for all $b\,\varepsilon\,K$. Since $b=as+sa\,\varepsilon\,K$, for $s\,\varepsilon\,S$, $a(as+sa)=0$, and so $aSa=(0)$; since $aKa=a^2K=(0)$, we obtain $aAa=(0)$, and so $a=0$ because A is a simple ring. Thus we must assume that every $a\neq 0\,\varepsilon\,K$ has an inverse in A. If $b\neq 0\,\varepsilon\,K$, $ab=\lambda\,\varepsilon\,Z$, $\lambda\neq 0$, whence $a\cdot ab=\lambda a$, and since $a^2=\mu\neq 0\,\varepsilon\,Z$, we are left with $b=\lambda' a$ where $\lambda'\,\varepsilon\,Z$. Thus $K=\{\lambda a\}$ for appropriate λ''s in Z.

If $s\,\varepsilon\,S$ then $(as-sa)^2=a\cdot sas+sas\cdot a-as^2a-sa^2a$, and since $sas\,\varepsilon\,K$, $a\cdot sas+sas\cdot a\,\varepsilon\,Z$; also $as^2a+sa^2s=as^2a+a^2s^2$ (since $a^2\,\varepsilon\,Z$) $=a(s^2a+as^2)$ $\varepsilon\,Z$ since $s^2a+as^2\,\varepsilon\,K$. Thus $(as-sa)^2\,\varepsilon\,Z$ for all $s\,\varepsilon\,S$. Since $S=[K,S]$

$+K \circ K$, and since $K = \{\lambda a\}$, and since $K \circ K \subset Z$, we have that every element s of S can be written as $s = \lambda(at - ta) + \mu$ for $\lambda, \mu \varepsilon Z$ and $t \varepsilon S$. Since every $r \varepsilon A$ can be written as $r = s + k$ we have that $r = \lambda(at + ta) + \mu + \alpha a$ for $\lambda, \mu, \alpha \varepsilon Z$, $t \varepsilon S$. Thus

$$(r - \mu)^2 = \lambda^2(at - ta)^2 + \alpha^2 a^2 + \alpha\lambda\{a(at - ta) + (at - ta)a\}.$$

Now, $(at - ta)^2 \varepsilon Z$, $a^2 \varepsilon Z$ and $a(at - ta) + (at - ta)a = a^2 t - ta^2 = 0$, so $(r - \mu)^2 \varepsilon Z$. Thus every $r \varepsilon A$ satisfies a quadratic equation over Z. As in the argument used in proving Theorem 9, A is at most 4-dimensional over Z. Theorem 15 is now completely proved.

It will be useful to show

LEMMA 16.[1] *If $u, v \varepsilon U$ then $uvu \varepsilon T(U)$ and $u^2 v + vu^2 \varepsilon T(U)$.*

Proof. To prove that $uvu \varepsilon T(U)$ we need but verify that $[uvu, K] \subset U$. Let $x \varepsilon K$. Thus

$$uvux - xuvu = \{u(vux + xuv) - (vux + xuv)u\} + \{vuxu - uxuv\}.$$

But $vux + xuv = vux - (vux)^* \varepsilon K$, so the first $\{\ \}$ is in U since U is a Lie ideal of K. Since $uxu \varepsilon K$ and $v \varepsilon U$, the second $\{\ \}$ is also in U. Thus $uvux - xuvu \varepsilon U$, and so $uvu \varepsilon T(U)$. Now

$$u^2 v + vu^2 = \{u(uv - vu) - (uv - vu)u\} + 2uvu,$$

and since $uvu \varepsilon T(U)$ and since $u(uv - vu) - (uv - vu)u \varepsilon U \subset T(U)$, $u^2 v + vu^2$ also is in $T(U)$, proving the lemma.

If U is a Lie ideal of K we define $B(U)$ by:

$$B(U) = \{x \varepsilon S \mid xu + ux \varepsilon T(U) \text{ for all } u \varepsilon U\}.$$

By Lemma 16, $u \varepsilon U$ implies that $u^2 \varepsilon B(U)$.

LEMMA 17. *Let $x \varepsilon B(U)$; then $(xu - ux)y + y(xu - ux) \varepsilon T(U)$ for all $y \varepsilon K$.*

Proof.

$$\begin{aligned}(xu - ux)y &+ y(xu - ux) \\ &= \{x(uy - yu) + (uy - yu)x\} + \{(xy + yx)u - u(xy + yx)\}.\end{aligned}$$

Since $uy - yu \varepsilon U$, and since $x \varepsilon B(U)$, the first $\{\ \}$ on the right-hand side is in $T(U)$. Since $xy + yx \varepsilon K$, and since $u \varepsilon U$, a Lie ideal of K, the second $\{\ \}$ is in U, and so is in $T(U)$; this proves the lemma.

[1] The proof is patterned after a suggestion of Willard E. Baxter.

The lemma motivates the following definition: for U a Lie ideal of K,

$$C(U) = \{x \varepsilon S \mid xy + yx \varepsilon T(U) \text{ for all } y \varepsilon K\}.$$

By Lemma 17, $[U, B(U)] \subset C(U)$.

Let $u, v \varepsilon U$. Thus $u^2 \varepsilon B(U)$. Hence $u^2v - vu^2 \varepsilon C(U)$. Therefore $(u^2v - vu^2)k + k(u^2v - vu^2) \varepsilon T(U)$ for all $k \varepsilon K$. On the other hand, if $s \varepsilon S$,

$$(u^2v - vu^2)s - s(u^2v - vu^2)$$
$$= \{u^2(vs - sv) - (vs - sv)u^2\} + \{(u^2s - su^2)v - v(u^2s - su^2)\}.$$

Since $vs - sv \varepsilon S$, the first $\{\ \}$ is in U by Lemma 10. Since $u^2s - su^2 \varepsilon K$, the second $\{\ \}$ is also in U. Thus the right-hand side is in U, consequently it is in $T(U)$. Thus $(u^2v - vu^2)s - s(u^2v - vu^2) \varepsilon T(U)$ for all $s \varepsilon S$. Given any element $r \varepsilon A$, $r = r_0 + r_1$ where $r_0 \varepsilon S$ and $r_1 \varepsilon K$. Computing $(u^2v - vu^2)r - ((u^2v - vu^2)r)^*$ in terms of r_0 and r_1 and using the above discussion we see that $(u^2v - vu^2)r - ((u^2v - vu^2)r)^*$ is in $T(U)$ for all $r \varepsilon A$. We summarize this in

THEOREM 18. *Let U be a Lie ideal of K, and let $u, v \varepsilon U$, $r \varepsilon A$; then*

$$(u^2v - vu^2)r - ((u^2v - vu^2)r)^* \varepsilon T(U).$$

We now define for U a Lie ideal of K, $G(U)$ by

$$G(U) = \{g \varepsilon A \mid gr - r^*g^* \varepsilon T(U) \text{ for all } r \varepsilon A\}.$$

By Theorem 18, $u^2v - vu^2 \varepsilon G(U)$ for all $u, v \varepsilon U$.

THEOREM 19. *Let A be a simple ring of characteristic $\neq 2$ and suppose that either Z, the center of A is (0) or that A is more than 4-dimensional over Z. If U is a Lie ideal of K and if $U \supset [K, K]$ then $G(U) = (0)$.*

Proof. Let $g \neq 0 \varepsilon G(U)$; thus for any $r \varepsilon A$

$$gr - r^*g^* \varepsilon T(U). \tag{1}$$

Thus, if $k \varepsilon K$, since $T(U)$ is a Lie ideal of K,

$$(gr - r^*g^*)k - k(gr - r^*g^*) \varepsilon T(U).$$

However,

$$(gr - r^*g^*)k - k(gr - r^*g^*) = g(rk) + (kr^*)g^* - r^*g^*k - kgr \varepsilon T(U).$$

Since $grk + kr^*g^* = g(rk) - (rk)^*g^*$, it is in $T(U)$ by (1). Thus we obtain

(2) $$r^*g^*k + kgr \varepsilon T(U) \text{ for all } r \varepsilon A,\ k \varepsilon K.$$

Let $k_1, k_2 \varepsilon K$. Since $T(U)$ is a Lie ideal of K, by (2) we have that

$$(r^*g^*k_1 + k_1gr)k_2 - k_2(r^*g^*k_1 + k_1gr) \varepsilon T(U).$$

That is,

$$k_1grk_2 - k_2r^*g^*k_1 + (rk_2)^*g^*k_1 - k_2k_1gr \varepsilon T(U).$$

However,

$$k_1grk_2 - k_2r^*g^*k_1 = k_1g(rk_2) + (rk_2)^*g^*k_1$$

and so is in $T(U)$ by (2). We thus are led to $r^*g^*k_1k_2 - (k_1k_2)^*gr \varepsilon T(U)$. Continuing in this way we obtain $r^*g^*\bar{k} - (\bar{k})^*gr \varepsilon T(U)$ for all $\bar{k} \varepsilon \bar{K}$ and all $r \varepsilon A$. By Theorem 15, $\bar{K} = A$ if the center of A is trivial or if A is more than 4-dimensional over Z, so in these cases we have that $r^*g^*t^* - tgr \varepsilon T(U)$ for all $t, r \varepsilon A$. Since A is simple, and since $g \neq 0$, $A = AgA$. Thus, given $y \varepsilon A$, $y = \sum r_igt_i$. Then $y^* = \sum t_i^*g^*r_i^*$, and so $y - y^* = \sum(r_igt_i - t_i^*g^*r_i^*)$. Since each $r_igt_i - t_i^*g^*r_i^* \varepsilon T(U)$ by the discussion above, we obtain that $y - y^* \varepsilon T(U)$ for all $y \varepsilon A$. However, every element in K has such a representation in the form $y - y^*$; we thus obtain $T(U) = K$. By the definition of $T(U)$ this is equivalent with $U \supset [K, K]$. Since we assumed $D \not\supset [K, K]$ we are led to a contradiction, and so $G(U) = (0)$.

We assume henceforth that $Z = (0)$ or that A is more than 4-dimensional over Z.

Since for U a Lie ideal of K, $u^2v - vu^2 \varepsilon G(U)$ for all $u, v \varepsilon U$, and since, by Theorem 19, if $U \not\supset [K, K]$, $G(U) = (0)$, we have

THEOREM 20. *If $U \not\supset [K, K]$ is a Lie ideal of K, and if $u, v \varepsilon U$, then $u^2v = vu^2$.*

We now prove

THEOREM 21. *Suppose U is a Lie ideal of K in which $u^2 = 0$ for every $u \varepsilon U$. Then $U = (0)$.*

Proof. We suppose that $u^2 = 0$ for every $u \varepsilon U$. Let $u \varepsilon U$, $k \varepsilon K$. Thus $2uku = (uk - ku)u - u(uk - ku) \varepsilon U$. Since $2K = K$, we obtain $uku \varepsilon U$ for all $k \varepsilon K$ and all $u \varepsilon U$. If $u, v \varepsilon U$ then $uv + vu = (u + v)^2 - u^2 - v^2 = 0$, so left-multiplying this by v we obtain $vuv = 0$ for all $v, u \varepsilon U$. Since $uku \varepsilon U$, we also have $v(uku)v = 0$. That is, $vukuv = 0$ for all $u, v \varepsilon U$ and $k \varepsilon K$; since $uv = -vu$ we can obtain $vukvu = 0$ for all $v, u \varepsilon U$ and all $k \varepsilon K$. Let $w \varepsilon U$ and $s \varepsilon S$. Thus $sw + ws \varepsilon K$, whence $vu(sw + ws)vu = 0$. Left-multiplying this by w we obtain $wvu(sw + ws)vu = 0$. However, since

$w \varepsilon U$, $wvu = -vwu = vuw$, thus $wvuwsvu = vuw^2svu = 0$, and since $0 = wvu(sw + ws)vu$, this reduces to $(wvu)s(wvu) = 0$; in other form this says $wvuSwvu = (0)$. Since we have already established that $wvuKwvu = (0)$, and since $A = K + S$, these combine to yield $wvuAwvu = (0)$. The simplicity of A then leads to $wvu = 0$ for all $w, v, u \varepsilon U$. If $k \varepsilon K$, let $w = ku - uk \varepsilon U$; hence $(ku - uk)vu = 0$, and since $uvu = 0$, we arrive at $ukvu = 0$ for all $k \varepsilon K$. Put $k = sv + vs$ where $s \varepsilon S$. Then $0 = u(sv + vs)vu = uvsvu = uvsuv$ since $v^2 = 0$, $uv = -vu$. That is, $uvSuv = (0)$. As we already have established that $uvKuv = (0)$, we reach $uvAuv = (0)$, and so $uv = 0$ for all $u, v \varepsilon U$. Put $v = uk - ku$ where $k \varepsilon K$. Since $u^2 = 0$, $uv = 0$ yields that $uku = 0$ for all $k \varepsilon K$. If $s \varepsilon S$, $sus \varepsilon K$ and so $u(sus)u = 0$, by the above. If $r \varepsilon A$, then $r = s + k$ when $s \varepsilon S$, $k \varepsilon K$ and so $ururu = u(s + k)u(s + k)u = ususu = 0$ since $uku = 0$. Thus uA is a nil right-ideal of A in which each element is nilpotent of index 3. As we have previously shown in this paper this is impossible in a simple ring unless $uA = (0)$; but then $u = 0$. Thus we have shown that $U = (0)$, proving Theorem 21.

THEOREM 22. *Let U be a Lie ideal of K and suppose that $U \not\supset [K, K]$. Then $u \varepsilon U$ implies that $u^2 \varepsilon Z$, the center of A.*

Proof. Since $u^2v = vu^2$ for all u, v in U, by Theorem 20, u^2 is in the center of $\bar{U}$, the subring generated by U. Now, $u^2s - su^2 \varepsilon U \subset \bar{U}$ by Lemma 10 for all $s \varepsilon S$. Also $u^2k - ku^2 = (uk - ku)u + u(uk - ku)$ and since $u \varepsilon U$, $uk - ku \varepsilon U$, each of $u(uk - ku)$ and $(uk - ku)u$ is in $\bar{U}$ for $k \varepsilon K$; thus $u^2k - ku^2 \varepsilon \bar{U}$. But then $u^2a - au^2 \varepsilon \bar{U}$ for all $a \varepsilon A$. Since u^2 is in the center of $\bar{U}$, $u^2(u^2a - au^2) = (u^2a - au^2)u^2$ for all $a \varepsilon A$, $u \varepsilon U$. The theorem will thus be proved when we prove

SUBLEMMA. *Let A be a simple ring of characteristic $\neq 2$. Suppose $t \varepsilon A$ is such that $t(ta - at) = (ta - at)t$ for all $a \varepsilon A$. Then $t \varepsilon Z$.*

This sublemma has some independent interest. To prove the sublemma we proceed as follows.

We know that $t(tr - rt) = (tr - rt)t$ for all $r \varepsilon A$. Let $p \varepsilon A$. Thus $t(trp - rpt) = (trp - rpt)t$. But $trp - rpt = (tr - rt)p + r(tp - pt)$. Thus $t(trp - rpt) = t\{(tr - rt)p + r(tp - pt)\} = (tr - rt)tp + tr(tp - pt)$ since $tr - rt$ commutes with t. Similarly $(trp - rpt)t = (tr - rt)pt + rt(tp - pt)$. Equating the two, transposing and simplifying we arrive at $2(tr - rt)(tp - pt) = 0$, and since the characteristic of A is not 2, we have that $(tr - rt)(tp - pt) = 0$ for all $r, p \varepsilon A$. In particular, for $p = ar$,

(1) $$(tr - rt)(tar - art) = 0 \qquad \text{for all } a, r \varepsilon A.$$

Since $tar - art = (ta - at)r + a(tr - rt)$ and since $(tr - rt)(ta - at) = 0$, (1) yields that $(tr - rt)a(tr - rt) = 0$; that is, $(tr - rt)A(tr - rt) = (0)$. This is impossible in a simple ring unless $tr - rt = 0$. Thus $t \varepsilon Z$ and the sublemma is established. Since $u^2(u^2a - au^2) = (u^2a - au^2)u^2$ in the theorem, $u^2 \varepsilon Z$ follows, and the theorem is proved.

COROLLARY. *Let U be a Lie ideal of K and suppose that $U \not\supset [K,K]$. Then $u, v \varepsilon U$ implies that $uv + vu \varepsilon Z$.*

We now are able to dispose of the situation in which $Z = (0)$. Indeed

THEOREM 23. *Let A be a simple ring of characteristic $\neq 2$ whose center $Z = (0)$. If $U \neq (0)$ is a Lie ideal of K then $U \supset [K,K]$.*

Proof. If $U \not\supset [K,K]$ then by Theorem 22 $u \varepsilon U$ implies that $u^2 \varepsilon Z = (0)$. Consequently $u^2 = 0$ for all $u \varepsilon U$. By Theorem 21 this results in $U = (0)$.

Having, in this way, settled the case $Z = (0)$ *we henceforth assume that* $Z \neq (0)$.

The $*$ of A induces an automorphism on Z. Two possibilities now confront us, namely

(1) $\lambda^* = \lambda$ for all $\lambda \varepsilon Z$, an involution of the first kind,

(2) $\mu^* \neq \mu$ for some $\mu \varepsilon Z$, an involution of the second kind.

In the case of an involution of the second kind, $\lambda = \mu^* - \mu \neq 0$ is in Z and is such that $\lambda^* = -\lambda$. Thus (2) is equivalent to

(2′) $\mu^* = -\mu$ for some $\mu \neq 0 \varepsilon Z$.

Our discussion first turns to the case (2′) in which there is a skew element in the center of A.

Let $\mu \varepsilon Z$, $\mu^* = -\mu \neq 0$. Let U be a Lie ideal of K and we further suppose that $U \not\supset [K,K]$.

If $s \varepsilon S$ then $\mu s \varepsilon K$, hence for all $u \varepsilon U$, $u(\mu s) - (\mu s)u \varepsilon U$; that is $\mu(us - su) \varepsilon U$. If $v \varepsilon U$, by the corollary to Theorem 22,

$$\mu(us - su)v + v(\mu(us - su)) \varepsilon Z. \text{ That is } \mu((us - su)v + v(us - su)) \varepsilon Z.$$

As a consequence, since $\mu \neq 0 \varepsilon Z$, $(us - su)v + v(us - su) \varepsilon U$ for all $s \varepsilon S$ and all $u, v \varepsilon U$. In particular, if $k \varepsilon K$, $k^2 \varepsilon S$, whence

$$z = (uk^2 - k^2u)v + v(uk^2 - k^2u)$$

is in Z. Since $uk^2 - k^2u = (uk - ku)k + k(uk - ku)$, it is readily verified that

discussion $(uk' - k'u)k(uk - ku) = 0$, we are left with $(uk - ku)K(uk - ku) = (0)$. If $s \varepsilon S$, $s(uk - ku)s \varepsilon K$, and so $(uk - ku)s(uk - ku)s(uk - ku) = 0$. If $r \varepsilon A$, $r = s + k'$ where $s \varepsilon S$ and $k' \varepsilon K$. By the above discussion it follows that $(uk - ku)(s + k')(uk - ku)(s + k')(uk - ku) = 0$. So $(uk - ku)A$ is a right-ideal of A each of whose elements is nilpotent of index of nilpotence 3, and as we have seen before, this forces $uk - ku = 0$ for all $k \varepsilon K$. Since $\bar{K} = A$ this places u in Z, violating $Z \cap K = (0)$. Thus for some $v \varepsilon U$ with $uv + vu = 0$ it must be that $v^2 = \lambda_2 \neq 0 \varepsilon Z$. This proves our contention.

So, we now have elements $u, v \varepsilon U$ such that $u^2 = \lambda_1 \neq 0$, $v^2 = \lambda_2 \neq 0$, $\lambda_1, \lambda_2, \varepsilon Z$ and $uv + vu = 0$. A simple calculation verifies that $(uv - vu)^2 = -4\lambda_1\lambda_2 \neq 0$.

We claim that u, v and $uv - vu$ are linearly independent over Z. For if $w = \lambda_0 u + \lambda_1 u + \lambda_2(uv - vu) = 0$, $\lambda_i \varepsilon Z$, $uw + wu = 0$ yields $2\lambda_0 = 0$, $vw + wv = 0$ yields $2\lambda_1 = 0$ and $(uv - vu)w + w(uv - vu) = 0$ yields that $2\lambda_2 = 0$. Hence $\lambda_0 = \lambda_1 = \lambda_2 = 0$, and so u, v, $uv - vu$ are linearly independent over Z.

Suppose now that we could find an $x \varepsilon U$ so that u, v, $uv - vu$ and x are linearly independent over Z. Suppose that $ux + xu = \lambda$, $vx + xv = \mu$ where $\lambda, \mu \varepsilon Z$. Consider $x' = x + \alpha u + \beta v$ where α and β are in Z. Now, $x'u + ux' = \lambda + 2\alpha\lambda_1$, $x'v + vx' = \mu + 2\beta\lambda_2$. Since $\lambda_1 \neq 0$, $\lambda_2 \neq 0$ we can solve these for α, β to force $x'u + ux' = x'v + vx' = 0$. Note that x' is linearly independent of u, v and $uv - vu$ and that it is in U, (since $Z \subset S$). We drop the $'$. Since $xu + ux = xv + vx = 0$, we have that $x(uv - vu) = (uv - vu)x$. Since $x \varepsilon U$, $uv - vu \varepsilon U$, then $x(uv - vu) + (uv - vu)x \varepsilon Z$, thus $2x(uv - vu) \varepsilon Z$. Since $x \neq 0$ and since $uv - vu$ has an inverse, $2x(uv - vu) = \lambda \neq 0 \varepsilon Z$. Multiplying both sides by $uv - vu$, and using that $(uv - vu)^2 = -4\lambda_1\lambda_2$, we obtain that $x = \lambda'(uv - vu)$, $\lambda' \varepsilon Z$, contradicting that x was linearly independent over Z of u, v and $uv - vu$. Thus no x linearly independent over Z with u, v, $uv - vu$ can be found in U. Thus *U is 3-dimensional over Z.* Also, to be more exact, U has a basis over Z consisting of u, v and $uv - vu$ where $u^2 = \lambda_1 \neq 0$, $v^2 = \lambda_2 \neq 0$, $\lambda_1, \lambda_2 \varepsilon Z$ and where $uv + vu = 0$.

Let $N(u) = \{x \varepsilon K \mid xu = ux\}$, and we similarly define $N(v)$. By the nature of the basis of U, it is clear that $W = N(u) \cap N(v)$ is such that $W = \{x \varepsilon K \mid [x, U] = (0)\}$.

We claim that W is a Lie ideal of K; for $s \varepsilon W$, $k \varepsilon K$, $t \varepsilon U$ implies that

$$(xk - kx)t - t(xk - kx)$$
$$= \{x(kt - tk) - (kt - tk)x\} + \{(xt - tx)k - k(xt - tx)\},$$

since $st - ts = 0$ because $t \varepsilon U$, the second $\{\ \}$ is 0; since $kt - tk \varepsilon U$, the first $\{\ \}$ is also 0. Thus $xk - kx \varepsilon W$. Also $W \not\supset [K,K]$ for $uv - vu \varepsilon [K,K]$ but $uv - vu \not\varepsilon W$. We should like to show that $W \neq (0)$.

If $s \varepsilon S$, u and v as previously chosen in U then we claim that $(us - su)v + v(us - su) \varepsilon W$. This can be verified readily by noting that $(us - su)v + v(us - su) = -\{u(vs - sv) + (vs - sv)u\}$ and that $u^2 \varepsilon Z$ and $v^2 \varepsilon Z$. So, if $W = (0)$ then $(us - su)v + v(us - su) = 0$ for all $s \varepsilon S$. But then, since in particular $vs - sv \varepsilon S$ when $s \varepsilon S$.

$$\{u(vs - sv) - (vs - sv)u\}v + v\{u(vs - sv) - (vs - sv)u\} = 0.$$

Since v commutes with $u(vs - sv)$ and with $(vs - sv)u$, the equation reduces to $2v(u(vs - sv) - (vs - sv)u) = 0$. Since v has an inverse in A we obtain that $u(vs - sv) - (vs - sv)u = 0$ for all $s \varepsilon S$. But $u(vs - sv) + (vs - vs)u = 0$ since it is in W. Thus $u(vs - sv) = 0$, and since u has an inverse in A, $vs - sv = 0$ for all $s \varepsilon S$. Therefore v is in the center of $\bar{S}$, the subring generated by S. Because A is more than 4-dimensional over Z, by Theorem 9 $\bar{S} = A$. Thus $x \varepsilon Z$; that is $v \varepsilon Z \cap K = (0)$, forcing $v = 0$, a contradiction. In this way we have proved that $W \neq (0)$.

Since $W \not\supset [K,K]$ and $W \neq (0)$ and it is a Lie ideal of K, W must also be 3-dimensional over Z. Consider $[W,K]$. Since W is a Lie ideal of K, $[W,K] \subset W$, and is also a Lie ideal of K. It can not be that $[W,K] = (0)$ for then $[W,\bar{K}] = (0)$, and since $\bar{K} = A$, $[W,A] = (0)$ which would imply that $W \subset Z \cap K = (0)$. Thus $[W,K]$ must also be 3-dimensional over Z; since it is contained in the 3-dimensional W, it follows that $[W,K] = W$. Thus $W \subset [K,K]$. Similarly $U \subset [K,K]$. As is seen by examining U, $U \cap W = (0)$. Now $W + U$ is a Lie ideal of K and is contained in $[K,K]$. It is 6-dimensional over Z. If $W + U \neq [K,K]$ then our previous discussion shows that its dimension over Z would be 3, which it is not. Thus $W + U = [K,K]$.

Suppose now that $[K,K] \neq K$. We claim that there exists an $s \varepsilon S$ so that $us + su \not\varepsilon [K,K]$. For, if $us + su$ is in $[K,K]$ for all $s \varepsilon S$, since $uk - ku \varepsilon [K,K]$, for all $k \varepsilon K$, we would have that $ua - a^*u^* \varepsilon [K,K]$ for all $a \varepsilon A$. As in the proof of Theorem 19 this leads to $[K,K] = K$. Thus, $us_0 + s_0u \not\varepsilon [K,K]$ for some $s_0 \varepsilon S$. Let $x_1 = us_0 + s_0u \varepsilon K$, $x_1 \not\varepsilon [K,K]$. Now $vx_1 - x_1v \varepsilon U$, so $vx_1 - x_1v = \alpha v + \beta u + \gamma(uv - vu)$ where $\alpha, \beta, \gamma \varepsilon Z$, since u, v and $uv - vu$ form a basis of U over Z. Since $(vx_1 - x_1v)v + v(vx_1 - x_1v) = 0$, we see that $\alpha = 0$. If $\beta = 0$, then $v(x_1 - \gamma u) = (x_1 - \gamma u)v$, and so $x_1 - \gamma u \varepsilon N(u) \cap N(v) = W \subset [K,K]$, from which $x_1 \varepsilon [K,K]$, a contra-

diction. If, on the other hand, $\beta \neq 0$, $v\left(\frac{x_1 - \gamma u}{\beta}\right) - \left(\frac{x_1 - \gamma u}{\beta}\right)v = u$, that is, there is a $t \varepsilon K$, $t \not\varepsilon [K,K]$ with $ut = tu$ and $u = vt - tv$. Now

$$(uv - vu)t - t(uv - vu)$$
$$= \{u(vt - tv) - (vt - tv)u\} + \{(ut - ut)v - v(ut - tu)\} = 0$$

since $ut - tu = 0$ and since $vt - tv = u$. Thus t commutes with $uv - vu = 2uv$ and with u. So, $utv = tuv = uvt$, thus since u has an inverse in A, $tv = vt$. But then $t \varepsilon N(u) \cap N(v) = W \subset [K,K]$, a contradiction. Thus the assumption that $[K,K] \neq K$ leads to a contradiction.

So we assume that $[K,K] = K$. Therefore K is 6-dimensional over Z.

Suppose that $s_1, s_2, \cdots, s_5$ in S are linearly independent over Z. Now $s_iu + us_i \varepsilon K$ and commute with u for $i = 1, 2, \cdots, 5$. Thus it is easily seen that $us_i + s_iu = \lambda_{i1}u + \lambda_{i2}w_1 + \lambda_{i3}w_2 + \lambda_{i4}w_3$ where the w_i's are a basis of W and where the $\lambda_{ij} \varepsilon Z$ for $i = 1, 2, \cdots, 5$. Therefore, for some $\alpha_i \varepsilon Z$, not all 0, $\sum_{i=1}^{5} \alpha_i(s_iu + us_i) = 0$. That is, $(\sum \alpha_i s_i)u + u(\sum \alpha_i s_i) = 0$, and $t = \sum \alpha_i s_i \neq 0$ since the s_i are linearly independent over Z. t is of course in S. Thus, given any 5 linearly independent elements in S we can produce from them an element $t \neq 0 \varepsilon S$ so that $ut + tu = 0$.

If the dimension of S over Z is larger than 124, we can find 25 groups of independent elements of S each group consisting of 5 members. From each group we get an element t_i, $i = 1, 2, \cdots, 25$ in S with $ut_i + t_iu = 0$ and where the t_i are linearly independent over Z. We split the t_i's into 5 groups of 5 elements in each. As with u, we obtain 5 elements $p_i \varepsilon S$, so that $p_iv + vp_i = 0$ for $i = 1, 2, \cdots, 5$, and where the p_i are linearly independent over Z. Since the p_i's are linear combinations of the t_i's, $p_iu + up_i = 0$ for $i = 1, 2, \cdots, 5$. Thus $p_iuv = uvp_i$, $p_ivu = vup_i$, and so $p_i(uv - vu) = (uv - vu)p_i$. However, as we did with u and v, since there are 5 linearly independent p_i's in Z, we can find an element $q \neq 0 \varepsilon S$ which is a linear combination of the p_i's for which $q(uv - vu) + (uv - vu)q = 0$. However, since q is a linear combination of the p_i's and since each p_i commutes with $uv - vu$, we have that $q(uv - vu) = (uv - vu)q$. Therefore $2q(uv - vu) = 0$. Since $2(uv - vu)$ has an inverse in A, $q = 0$ must follow, contradicting that $q \neq 0$. Thus we must assume that the dimension of S over Z is less than 125. But then, since $A = S + K$, the dimension of A over Z is at most 131. By the known results for finite-dimensional simple algebras [7, 8], if A is not the 4×4 matrices over a field Z, U must contain $[K,K]$.

We have finally proved

THEOREM 25. *Let A be a simple ring of characteristic $\neq 2$, and suppose that A is more than 16-dimensional over its center $Z \neq (0)$. Suppose further that $\lambda \varepsilon Z$ implies that $\lambda^* = \lambda$. If $U \neq (0)$ is a Lie ideal of K then U must contain $[K, K]$.*

Combining Theorems 23, 24 and 25 we obtain the main theorem of this section, namely

THEOREM 26. *Let A be a simple ring of characteristic $\neq 2$, with an involution and suppose that either $Z = (0)$ or that A is more than 16-dimensional over Z, its center; if K is the set of skew elements of A then every Lie ideal, U, of K must satisfy*

(1) *either* $U \subset Z$ *or* (2) $U \supset [K, K]$.

We combined Theorems 23, 24 and 25 to get the general Theorem 26; however, the sum total of information contained in these three theorems separately exceeds that given in the statement of Theorem 26.

We close the paper with

THEOREM 27. *Let A be as in Theorem 26. Then*

$$[[K, K], [K, K]] = [K, K].$$

Proof. Let $U = [[K, K], [K, K]]$. U is certainly a Lie ideal of K. Thus if $U \neq [K, K]$, then it is strictly contained in $[K, K]$, therefore by Theorem 26, $U \subset Z$. Let $a \varepsilon [K, K]$, $k \varepsilon K$. Thus

$$b = (ak - ka)a - a(ak - ka) \varepsilon U \subset Z.$$

Since aka is in K along with k, if we replace k by aka in the expression for b, and simplify the expression, we have that $aba \varepsilon Z$. Since $b \varepsilon Z$, $aba = a^2 b \varepsilon Z$. If $b \neq 0$ it has an inverse in Z, and so $a^2 \varepsilon Z$. If $a^2 \not\varepsilon Z$, then $b = 0$ for all $k \varepsilon K$, so $a(ak - ka) = (ak - ka)a$ for all $k \varepsilon K$. That is,

$$(1) \qquad a^2 k + k a^2 = 2aka \qquad \text{for all } k \varepsilon K.$$

Consider $k = as + sa \varepsilon K$ where $s \varepsilon S$. For this k (1) yields

$$0 = a^2 k + k a^2 - 2aka = a^2(as - sa) - (as - sa)a^2,$$

thus

$$(2) \qquad a^2(as - sa) = (as - sa)a^2 \qquad \text{for all } s \varepsilon S.$$

Since $a^2(ak - ka) = (ak - ka)a^2$, we obtain $a^2(ar - ra) = (ar - ra)a^2$ for all $r \varepsilon A$. Now, since $a^2 r - r a^2 = a(ar - ra) + (ar - ra)a$ we have that $a^2(a^2 r - r a^2) = (a^2 r - r a^2)a^2$ for all $r \varepsilon A$. By the sublemma of Theorem 22,

$a^2 \varepsilon Z$. Thus in any case $a^2 \varepsilon Z$ for all $a \varepsilon [K,K]$. If $a^2 = 0$ for $a \varepsilon [K,K]$, then since $a(ak - ka) - (ak - ka)a = -2aka \varepsilon U \subset Z$, we obtain $aka \varepsilon Z$, and since a has no inverse in A, $aka = 0$ for all $k \varepsilon K$. As we have shown several times earlier, this forces $a = 0$.

Consequently $a \neq 0 \varepsilon [K,K]$ implies that $a^2 = \lambda \neq 0 \varepsilon Z$.

Now $Z \ni a(ak - ka) - (ak - ka)a = 2a(ak - ka)$, whence $a(ak - ka) \varepsilon Z$ for all $a \varepsilon [K,K]$ and all $k \varepsilon K$. Since a has an inverse in A, this implies that $a(ak - ka) = (ak - ka)a$ for all $k \varepsilon K$. However, since $a^2 \varepsilon Z$, $a(ak - ka) + (ak - ka)a = 0$, and so $2a(ak - ka) = 0$. Since a is regular, this forces $ak - ka = 0$ for all $k \varepsilon K$. Thus a is in the center of $\bar{K} = A$. That is, $[K,K] \subset Z$.

Repeating the argument used above on K instead of $[K,K]$ this time, we arrive at $K \subset Z$. But then $[K,K] = (0)$ and so $[[K,K],[K,K]] = [K,K] = (0)$, and the theorem is proved.

UNIVERSITY OF PENNSYLVANIA.

REFERENCES.

[1] Willard E. Baxter, "Lie simplicity of a special class of associative ring," to appear.

[2] J. Dieudonné, "On the structure of unitary groups," *Transactions of the American Mathematical Society*, vol. 72 (1952), pp. 367-385.

[3] I. N. Herstein, "On the Lie and Jordan rings of a simple associative ring," *American Journal of Mathematics*, vol. 77 (1955), pp. 279-285.

[4] ———, "On the Lie ring of a division ring," *Annals of Mathematics*, vol. 60 (1954), pp. 571-575.

[5] ———, "On the Lie ring of a simple ring," *Duke Mathematical Journal*, vol. 22 (1955), pp. 471-476.

[6] N. Jacobson, "Structure theory for algebraic algebras of bounded degree," *Annals of Mathematics*, vol. 46 (1945), pp. 695-707.

[7] ———, "Classes of restricted Lie algebras of characteristic p. I," *American Journal of Mathematics*, vol. 63 (1941), pp. 481-515.

[8] I. Kaplansky, "Report on seminar on simple Lie algebras," *Bulletin of the American Mathematical Society*, vol. 60 (1954), pp. 470-471.

[9] J. Levitski, "A problem of A. Kurosh," *Bulletin of the American Mathematical Society*, vol. 52 (1946), pp. 1033-1035.

[10] ———, "On the radical of a general ring," *ibid.*, vol. 49 (1943), pp. 462-466.

PARTIAL DIFFERENCE SETS.*

By D. R. Hughes.[1]

1. Introduction. The notion of a transitive projective plane and the resulting characterization of the plane by a group with a difference set have been the subject of much interest in recent years (see [5, 9]). In [10] a somewhat similar situation, a transitive affine plane, is studied, and in [11] another similar situation arose in the investigation of associative planar division neo-rings. In this paper we introduce a generalization of the above situations which includes quite a number of other types of projective planes. In each case the plane is characterized by a (collineation) group with a particular subgroup structure and a subset called a "partial difference set."

2. Planar ternary rings. Some use will be made of the planar ternary rings developed by Hall ([8]), but with a different coordinatizing scheme and a different notation; in particular, the ternary function F and the coordinatizing scheme of [11] will be used. In [8] Hall has developed a number of equivalences between the algebraic structure of the planar ternary ring and the geometric structure of the plane, and in [13] more equivalences of this type will be found (indeed, [13] is a very good source for all of these equivalences). All of these results, with perhaps slight modifications to account for the different coordinatizing scheme, carry over to the scheme used here, and those that will be needed will be listed. First a brief sketch of the ternary ring will be given.

The *planar ternary ring* (R, F) is a set R containing at least the two distinct elements 0 (zero) and 1 (one), together with a *ternary function* F (mapping ordered triples of R upon R) satisfying:

(A) $F(a,0,c) = F(0,b,c) = c$, all $a, b, c \in R$;

(B) $F(a,1,0) = F(1,a,0) = a$, all $a \in R$;

* Received September 14, 1955; revised, April 20, 1956.

[1] This research was carried out while the author was a National Science Foundation Fellow.

(C) if $a, b, c, d \in R$, $a \neq c$, then there is a unique $x \in R$ such that $F(x, a, b) = F(x, c, d)$;

(D) if $a, b, c \in R$, then there is a unique $x \in R$ such that $F(a, b, x) = c$;

(E) if $a, b, c, d \in R$, $a \neq c$, then there is a unique ordered pair $x, y \in R$ such that $F(a, x, y) = b$, $F(c, x, y) = d$.

For all $a, b \in R$, $a \cdot b$ or ab is defined to be $F(a, b, 0)$, and $a + b$ is defined to be $F(1, a, b)$. Then the set R^* of non-zero elements of R is a loop under the operation $(\cdot)$ with identity 1, and R is a loop under the operation $(+)$ with identity 0; these are the multiplicative and additive loops, respectively, of (R, F). If $F(a, b, c) = ab + c$ for all $a, b, c \in R$ then the ring is said to be *linear*.

In the coordinate scheme of [11] the plane π is related to (R, F) as follows: points are (a, b), (a), (∞), for all $a, b \in R$; lines are $[m, k]$, $[\infty, (k, 0)]$, L_∞, for all $m, k \in R$. The rules of incidence are: (a, b) is on $[m, k]$ if $F(m, a, b) = k$, and (a, b) is on $[\infty, (a, 0)]$; (m) is on $[m, k]$ and L_∞; (∞) is on $[\infty, (k, 0)]$ and L_∞. Note that this scheme is different from the scheme in [13], as well as that in [8].

In any projective plane if P is a point and L is a line, then the plane is said to be (P, L) *transitive* (*with group* H) if there is a group H of collineations which fixes every point on L and every line through P and which is transitive and regular on the "non-fixed" points on any line through P. That is, if Q and R are points collinear with P, both distinct from P and neither on L, then there is a unique $h \in H$ such that $Qh = R$. For further details about (P, L) transitivity, see [2, 13]; we remark that a plane is (P, L) transitive if and only if Desargues' Theorem is valid with P as the center of perspectivity and L as the line of perspectivity.

Now assume that π is coordinatized by the planar ternary ring (R, F) with $X = (0)$, $Y = (\infty)$, $O = (0, 0)$.

THEOREM 1. *(R, F) is linear with associative addition if and only if π is (Y, XY) transitive with group H; furthermore, the additive loop of (R, F) is then isomorphic to H.*

THEOREM 2. *(R, F) is linear with associative multiplication if and only if π is (X, OY) transitive with group H; furthermore, the multiplicative loop of (R, F) is then isomorphic to H.*

We define the *left distributive law* to be $a(b + c) = ab + ac$, and say that (R, F) is *left distributive* if this law holds for all elements of R. The right distributive law is defined analogously, as is right distributivity.

THEOREM 3. *(R, F) is linear, left distributive, and has associative addition if and only if π is (P, XY) transitive for every point P on the line XY.*

THEOREM 4. *(R, F) is linear, left distributive (right distributive), and has associative multiplication if and only if π is (X, OY) and (Y, OX) transitive $((X, OY)$ and (O, XY) transitive).*

The proofs of the above theorems can be found in [13], or can be obtained by modifying or extending the proofs slightly. A planar ternary ring satisfying the hypotheses of Theorem 3 will be called a *left Veblen-Wedderburn system*, or merely a left V-W system (and a right V-W system is defined similarly), and a left V-W system with associative multiplication will be called a *left near-field.*

3. Partially transitive planes. Throughout the rest of this paper "projective plane" will always mean "finite projective plane." Then if π is a projective plane there will be an integer n (called the *order* of π) such that every point (line) is incident with $n+1$ lines (points) and such that π contains a total of n^2+n+1 points (lines).

Suppose π is a projective plane of order n, and $\mathfrak{G}$ is a non-trivial group of collineations of π; let π_0 be the set of points and lines of π that are fixed (element-wise) by every collineation of $\mathfrak{G}$. Let the points and lines of π_0 be called fixed points and fixed lines; let the points (lines) of π that are not in π_0 but are on lines (contain points) of π_0 be called tangent points (tangent lines); let the remaining points (lines) be called ordinary points (ordinary lines). Finally, suppose that $\mathfrak{G}$ is transitive and regular on both the ordinary points and the ordinary lines; i.e., if X, Y is an ordered pair of ordinary points (or lines) then there is a unique $g \in \mathfrak{G}$ such that $Xg = Y$. Then we shall say that π is a *partially transitive and regular plane with respect to $\mathfrak{G}$ and π_0*, or merely that π is *partially transitive.*

Obviously π_0 is either a degenerate subplane or a (non-degenerate) subplane of π (see [8] for a list of all degenerate planes). Since $\mathfrak{G}$ is transitive and regular on both ordinary points and ordinary lines, the number of ordinary points equals the number of ordinary lines (and this common number is the order of $\mathfrak{G}$); hence it is easy to see that π_0 contains the same number of points and lines. Thus π_0 must be one of the following types:[2]

[2] In the abstract of the paper "Partial difference sets," presented to the American Mathematical Society on Oct. 22, 1955, the author stated essentially that type $(4, m)$ only occurs if $m = 3$, and that types $(5, m)$, $(6, m)$ never occur. This is incorrect, and

(0) π_0 is empty.

(1a) π_0 consists of a point Q_0 and a line K_0, Q_0 on K_0.

(1b) π_0 consists of a point Q_0 and a line K_0, Q_0 not on K_0.

(2) π_0 consists of two points Q_0 and Q_1 and two lines K_0 and K_1, where $K_0 = Q_0Q_1$ and Q_0 is on K_1.

(3) π_0 consists of three non-collinear points Q_i, $i = 0, 1, 2$, and the three lines $K_0 = Q_1Q_2$, $K_1 = Q_0Q_2$, $K_2 = Q_0Q_1$.

$(4, m)$ π_0 consists of m points $(m \geqq 3)$ Q_i, $i = 1, 2, \cdots, m$, on a line K_0, a point Q_0 not on K_0, and the $m + 1$ lines K_0, $K_i = Q_0Q_i$.

$(5, m)$ π_0 consists of $m + 1$ points $(m \geqq 2)$ Q_i, $i = 0, 1, \cdots, m$ on a line K_0, and $m + 1$ lines K_i, $i = 0, 1, \cdots, m$, each through Q_0.

$(6, m)$ π_0 is a subplane (i.e., non-degenerate) of order m, with points Q_i, lines K_i, $i = 0, 1, \cdots, m^2 + m$.

Types (1a), (1b), (2), (3) would be special cases of types $(4, m)$ and $(5, m)$ if we allowed $m < 3$ or $m < 2$; however, the analysis is quite different in case $m \geqq 3$ or $m \geqq 2$, respectively, so we separate these cases. The results in this section will apply to all of the types, and we do not distinguish them until subsequent sections.

LEMMA 1. *Every tangent line contains ordinary points.*

Proof. Suppose, if possible, that L is a tangent line containing the fixed point Q and n tangent points $A_1, A_2, \cdots, A_n$. Through each point A_i there is a fixed line K_i and the intersection of two such fixed lines is a fixed point. If all the lines K_i contain the single point B then (since BQ is certainly a fixed line) every line through B is a fixed line, and so there are no ordinary points in the plane. This contradicts the assumption that $\mathfrak{G}$ is non-trivial. Note that if $n = 2$ then there is a point B on all these lines K_i, since there are only two such lines, So we can assume that all the K_i do not pass through a single point, and that $n > 2$. Hence it is easy to see that π_0 contains a set of four points, no three of them collinear, and thus π_0 is a non-degenerate subplane of π, of order m, say. If $n = m^2$, then every point of π is on a line of π_0, and this would contradict the assumption that $\mathfrak{G}$ is non-trivial. Thus (see, for instance, [5]) we know that $n \geqq m^2 + m$. Since π_0 contains at least the n lines K_i, and since Q is on at least 4 lines of π_0,

---

an example of type $(4, m)$ with $m = 4$ is given in Section 5; infinitely many examples of type $(5, m)$ are given in Section 6. The author has no examples of type $(6, m)$ but has no proof that they cannot occur.

π_0 must contain at least $n+4$ lines; so $m^2+m+1 \geqq n+4$. This contradicts $n \geqq m^2+m$, and the lemma is proven.

Now let P_0 be an arbitrarily chosen ordinary point, and let J_0 be an arbitrarily chosen ordinary line. Let D be the set of all $d \in \mathfrak{G}$ such that P_0d is on J_0. Let $R_i = J_0K_i$ and $L_i = P_0Q_i$; let $\mathfrak{R}_i$ be the subgroup of $\mathfrak{G}$ which fixes the point R_i, and let $\mathfrak{L}_i$ be the subgroup of $\mathfrak{G}$ which fixes the line L_i. Since $\mathfrak{R}_i$ consists exactly of those $x \in \mathfrak{G}$ such that J_0x contains R_i, $\mathfrak{R}_i$ has order equal to the number of ordinary lines through R_i; similarly $\mathfrak{L}_i$ has order equal to the number of ordinary points on L_i. By Lemma 1 each tangent line through Q_i contains an ordinary point and is thus an image L_ix of L_i; so it is fixed by the group $x^{-1}\mathfrak{L}_ix$. Similarly for the tangent points on K_i: each is fixed by a conjugate of $\mathfrak{R}_i$. Also, it is clear that $\mathfrak{R}_i \cap \mathfrak{R}_j = \mathfrak{L}_i \cap \mathfrak{L}_j = 1$ if $i \neq j$; for otherwise there would be a non-identity element of $\mathfrak{G}$ which would fix $P_0 = L_iL_j$, or $J_0 = R_iR_j$.

LEMMA 2. *If Q_i is not on K_j, then $\mathfrak{L}_i$ and $\mathfrak{R}_j$ are conjugates in $\mathfrak{G}$; thus if there is a line of π_0 which does not contain either Q_i or Q_j then $\mathfrak{L}_i$ and $\mathfrak{L}_j$ are conjugates in $\mathfrak{G}$.*

Proof. It is only necessary to prove the first part of the lemma. Suppose Q_i not on K_j; then the line $L = Q_iR_j$ is a tangent line, so $L = L_ix$ for some $x \in \mathfrak{G}$, by Lemma 1. Hence R_j is fixed by both $\mathfrak{R}_j$ and $x^{-1}\mathfrak{L}_ix$, so $\mathfrak{R}_j = x^{-1}\mathfrak{L}_ix$.

THEOREM 5. (a) *If $g \in \mathfrak{G}$, $g \notin \mathfrak{L}_i$ for any i, then $g = d_1d_2^{-1}$ for a unique ordered pair $d_1, d_2 \in D$; if $g \in \mathfrak{L}_i$ for some $i, g \neq 1$, then $g \neq d_1d_2^{-1}$ for any $d_1, d_2 \in D$.*

(b) *If $g \in \mathfrak{G}$, $g \notin \mathfrak{R}_i$ for any i, then $g = d_1^{-1}d_2$ for a unique ordered pair $d_1, d_2 \in D$; if $g \in \mathfrak{R}_i$ for some $i, g \neq 1$, then $g \neq d_1^{-1}d_2$ for any $d_1, d_2 \in D$.*

(c) $\mathfrak{R}_i \cap \mathfrak{R}_j = \mathfrak{L}_i \cap \mathfrak{L}_j = 1$ *if $i \neq j$.*

Proof. Let $g \in \mathfrak{G}, g \neq 1$, and consider the line $P_0 \cdot P_0g = L$. Either $L = J_0b$ for a unique $b \in \mathfrak{G}$, or $L = L_i$ for a unique i, and in the latter case (and only then) $g \in \mathfrak{L}_i$. If $L = J_0b$, then P_0b^{-1}, P_0gb^{-1} are both on J_0, so $b^{-1} = d_2 \in D$, $gb^{-1} = d_1 \in D$, or $g = d_1d_2^{-1}$. By a reversal of the above it is clear that d_1 and d_2 are unique and the cases of (a) are mutually exclusive. Similarly, (b) is proven using the point $J_0 \cdot J_0g$, and (c) has already been demonstrated.

LEMMA 3. *If $\mathfrak{G}$ contains an element b of order two, then $b \in \mathfrak{L}_i \cap \mathfrak{R}_j$ for some i and j.*

Proof. Suppose $b \notin \mathfrak{L}_i$ for any i. Then by Theorem 5, $b = d_1d_2^{-1}$ for a

unique ordered pair $d_1, d_2 \in D$. However, $b = b^{-1} = d_2 d_1^{-1}$, whence $d_1 = d_2$ and $b = 1$, a contradiction. Similarly, $b \notin \mathfrak{R}_i$ for any i is contradictory.

Of the various types given in this section, we shall henceforth exclude type (0), as that is the type given by Bruck in [5] (and including as an important special case the cyclic difference sets of Hall; see [9]). The remaining types will be analyzed in more detail, and we shall consistently use the terminology and notation introduced in this section. Noting Theorem 5, we shall refer to D as a *partial difference set* for the group $\mathfrak{G}$. Of course, a knowledge of the subgroups $\mathfrak{R}_i$ and $\mathfrak{L}_j$ is also needed for a complete description of the situation. Indeed, the existence of a group $\mathfrak{G}$ with a partial difference set D satisfying (a), (b), (c) of Theorem 5, together with certain conditions on the subgroups $\mathfrak{R}_i$ and $\mathfrak{L}_j$, implies the existence of a projective plane of the appropriate type. These other conditions vary somewhat for the different types, and will be given as they arise.

The following table is the result of straightforward counting, which we omit:

Type	Order of $\mathfrak{G}$	Order of $\mathfrak{R}_0$ and $\mathfrak{L}_0$	Order of $\mathfrak{R}_i$ and $\mathfrak{L}_i$, $i \neq 0$	Number of elements in D
(1a)	n^2	n		n
(1b)	$n^2 - 1$	$n - 1$		n
(2)	$n^2 - n$	n	$n - 1$	$n - 1$
(3)	$(n-1)^2$	$n - 1$	$n - 1$	$n - 2$
$(4, m)$	$(n-1)(n-m+1)$	$n - 1$	$n - m + 1$	$n - m$
$(5, m)$	$n(n-m)$	n	$n - m$	$n - m$
$(6, m)$	$(n-m)(n-m^2)$	$n - m^2$	$n - m^2$	$n - m^2 - m$

Note that types (3) and $(6, m)$ are "symmetric" in the sense that the subgroups for $i = 0$ are not particularly distinguished from those for $i \neq 0$.

4. Types (1a), (1b), (2), (3).

These types are somewhat simpler than the others, and generally admit of many examples. With the exception of type (1b), there are both Desarguesian and non-Desarguesian examples of each type known. Furthermore, quite a bit of simplification occurs if $\mathfrak{G}$ is abelian or if certain of the subgroups $\mathfrak{R}_i$ or $\mathfrak{L}_i$ are normal, a situation which cannot occur in the other types. We will first give some examples; in each case the example will be given in terms of a particular planar ternary ring of the plane under consideration.

Type (1a). Let (R, F) be a (not necessarily associative) division ring

(with identity). For each ordered pair $a, b \in R$, consider the mapping $\phi = \phi(a, b)$ defined below:

$$\begin{array}{ll} \phi\colon (x, y) \to (x + a, y + ax + b) & [m, k] \to [m - a, k + ma + b - a^2] \\ (m) \to (m - a) & [\infty, (k, 0)] \to [\infty, (k + a, 0)] \\ (\infty) \to (\infty) & L_\infty \to L_\infty. \end{array}$$

Then the set of all such mappings is a group $\mathfrak{G}$ of collineations with respect to which the plane is of type (1a), with $Q_0 = (\infty)$, $K_0 = L_\infty$. If we let $P_0 = (0, 0)$, $J_0 = [0, 0]$, then $\mathfrak{R}_0 = \mathfrak{L}_0 = \{\phi(0, b)\}$ is normal in $\mathfrak{G}$, and $D = \{\phi(a, 0)\}$.

Type (1b). Examples of this type will be found in [4].

Type (2). Let (R, F) be a linear planar ternary ring with associative addition and multiplication. For each ordered pair $a, b \in R$, $a \neq 0$, consider the mapping $\phi = \phi(a, b)$ defined below:

$$\begin{array}{ll} \phi\colon (x, y) \to (ax, y + b) & [m, k] \to [ma^{-1}, k + b] \\ (m) \to (ma^{-1}) & [\infty, (k, 0)] \to [\infty, (ak, 0)] \\ (\infty) \to (\infty) & L_\infty \to L_\infty. \end{array}$$

Then the set of all such mappings is a group $\mathfrak{G}$ of collineations with respect to which the plane is of type (2), with $Q_0 = (\infty)$, $Q_1 = (0)$, $K_0 = L_\infty$, $K_1 = [\infty, (0, 0)]$. If we let $P_0 = (1, 0)$, $J_0 = [1, 0]$, then $\mathfrak{R}_0 = \mathfrak{L}_0 = \{\phi(1, b)\}$ is normal in $\mathfrak{G}$, and $\mathfrak{R}_1 = \mathfrak{L}_1 = \{\phi(a, 0)\}$ is normal in $\mathfrak{G}$; D consists of all elements $\phi(a, -a)$.

Type (3). Let (R, F) be a linear planar ternary ring with the left distributive law and associative multiplication. For each ordered pair $a, b \in R$, $a \neq 0$, $b \neq 0$, consider the mapping $\phi = \phi(a, b)$ defined below:

$$\begin{array}{ll} \phi\colon (x, y) \to (ax, by) & [m, k] \to [bma^{-1}, bk] \\ (m) \to (bma^{-1}) & [\infty, (k, 0)] \to [\infty, (ak, 0)] \\ (\infty) \to (\infty) & L_\infty \to L_\infty. \end{array}$$

Then the set of all such mappings is a group $\mathfrak{G}$ of collineations with respect to which the plane is of type (3), with $Q_0 = (\infty)$, $Q_1 = (0)$, $Q_2 = (0, 0)$, $K_0 = [0, 0]$, $K_1 = [\infty, (0, 0)]$, $K_2 = L_\infty$. If we let $P_0 = (1, 1)$, $J_0 = [1, 1]$, and let $e \in R$ satisfy $e + 1 = 0$, then $\mathfrak{R}_0 = \mathfrak{L}_0 = \{\phi(1, b)\}$, $\mathfrak{R}_1 = \mathfrak{L}_1 = \{\phi(a, 1)\}$, $\mathfrak{R}_2 = \{\phi(a, a)\}$, $\mathfrak{L}_2 = \{\phi(a, eae^{-1})\}$. Each of these subgroups except the last two are normal in $\mathfrak{G}$; D consists of all $\phi(a, b)$ such that $a + b = 1$.

By an obvious modification, if "right distributive" is substituted for

"left distributive" in the above example of type (3), then the plane is still of type (3). Since there exist finite non-associative division rings (for type (1a)) and finite near-fields which are not fields (for types (2) and (3)), it is clear that not only does every Desarguesian plane yield an example of the above types but also that there are non-Desarguesian examples. The examples of type (1b) given in [4] are all Desarguesian, and $\mathfrak{G}$ is cyclic, and it is conjectured that all such finite cyclic examples are Desarguesian (see [10]); the author does not know of any finite non-cyclic examples.

From Lemma 2, the groups $\mathfrak{R}_0$ and $\mathfrak{L}_0$ are conjugate in type (1b); $\mathfrak{R}_1$ and $\mathfrak{L}_1$ are conjugate in type (2); for each i, $\mathfrak{R}_i$ and $\mathfrak{L}_i$ are conjugate in type (3). Furthermore, by a proper choice of P_0 and J_0 it is easy to see that certain of the $\mathfrak{R}_i$ and $\mathfrak{L}_i$ can be made to coincide: for $i=0$ in type (1b), $i=1$ in type (2), and any two values of i in type (3).

THEOREM 6. *If either $\mathfrak{R}_i$ or $\mathfrak{L}_i$ is normal in $\mathfrak{G}$, then $\mathfrak{R}_i=\mathfrak{L}_i$.*

Proof. As noted above, this is a corollary of Lemma 2 for every case except $i=0$ in types (1a) and (2), so we restrict attention to these cases. If $\mathfrak{R}_0$ is normal and $x\in\mathfrak{R}_0$, $x\neq 1$, then certainly $x\notin\mathfrak{L}_1$ (in type (2)), since the orders of $\mathfrak{R}_0$ and $\mathfrak{L}_1$ are relatively prime. If also $x\notin\mathfrak{L}_0$ then $x=d_1d_2^{-1}$ for some $d_1,d_2\in D$, whence $d_2^{-1}xd_2=d_2^{-1}d_1\in\mathfrak{R}_0$, since $\mathfrak{R}_0$ is normal. This is impossible, by (b) of Theorem 5, so we must have $x\in\mathfrak{L}_0$, whence $\mathfrak{R}_0=\mathfrak{L}_0$.

THEOREM 7. *In types* (1a) *and* (2), *if $\mathfrak{R}_0$ is normal in $\mathfrak{G}$, then π is coordinatizable by a linear planar ternary ring with associative addition.*

Proof. If $\mathfrak{R}_0$ is normal, then every tangent point R_0x on K_0 is fixed by $\mathfrak{R}_0=x^{-1}\mathfrak{R}_0x$, whence it is immediate that π is (Q_0,K_0) transitive. Thus the theorem follows from Theorem 1.

THEOREM 8. *In type* (1b) *if $\mathfrak{R}_0$ is normal in $\mathfrak{G}$, or in type* (2) *if $\mathfrak{R}_1$ is normal in $\mathfrak{G}$, then π is coordinatizable by a linear planar ternary ring with associative multiplication.*

Proof. Analogous to the proof of Theorem 7, using Theorem 2 instead of Theorem 1.

THEOREM 9. *In type* (2), *if both of the $\mathfrak{R}_i$ are normal in $\mathfrak{G}$ then π is coordinatizable by a linear planar ternary ring with associative addition and multiplication.*

Proof. Immediate from Theorems 7 and 8, noting that the same ternary ring can be used in the conclusion of each of these theorems.

Thus, referring to the example of type (2) given at the beginning of this section, Theorem 9 actually affords us a necessary and sufficient condition that a plane be coordinatizable by a linear ring with both operations associative.

THEOREM 10. *In type* (3), *if* $\mathfrak{R}_0$ *is normal in* $\mathfrak{G}$ *then* π *is coordinatizable by a linear planar ternary ring with associative multiplication. If* $\mathfrak{R}_0$ *and* $\mathfrak{R}_1$ *are both normal in* $\mathfrak{G}$ *then* π *is coordinatizable by a linear planar ternary ring with the left distributive law and associative multiplication; furthermore,* $\mathfrak{G}$ *is isomorphic to the direct product of* $\mathfrak{R}_0$ *with itself, and all three of the* $\mathfrak{R}_i$ *are isomorphic to one another. If all three of the* $\mathfrak{R}_i$ *are normal in* $\mathfrak{G}$ *then* π *is coordinatizable by a linear planar ternary ring with both distributive laws and associative, commutative multiplication* (*i. e., an abelian planar division neo-ring; see* [11]).

Proof. The first sentence and the first part of the second sentence in the theorem are immediate from Theorems 2 and 4. Assume that $\mathfrak{R}_0$ and $\mathfrak{R}_1$ are normal in $\mathfrak{G}$; it is clear that $\mathfrak{G}$ is their direct product. If $x \in \mathfrak{R}_2$, then $x = ab$, $a \in \mathfrak{R}_0$, $b \in \mathfrak{R}_1$, and this representation is unique; furthermore, for any $a \in \mathfrak{R}_0$, there is exactly one $b \in \mathfrak{R}_1$ such that $ab \in \mathfrak{R}_2$, since $\mathfrak{R}_0 \cap \mathfrak{R}_2 = \mathfrak{R}_1 \cap \mathfrak{R}_2 = 1$. Thus every element of $\mathfrak{R}_2$ can be written in the form $a(aT)$ for some $a \in \mathfrak{R}_0$, where T is a one-to-one mapping of $\mathfrak{R}_0$ upon $\mathfrak{R}_1$. If $a, b \in \mathfrak{R}_0$, then $[a(aT)][b(bT)] = ab[(aT)(bT)] \in \mathfrak{R}_2$, so $(aT)(bT) = (ab)T$. Hence $\mathfrak{R}_0$ is isomorphic to $\mathfrak{R}_1$, and obviously $\mathfrak{R}_2 = \{a(aT)\}$ is also isomorphic to $\mathfrak{R}_0$.

If also $\mathfrak{R}_2$ is normal, let $a, b \in \mathfrak{R}_0$. Then $a^{-1}[b(bT)]a = a^{-1}ba(bT) \in \mathfrak{R}_2$, so $bT = (a^{-1}ba)T$ or $b = a^{-1}ba$. Thus $\mathfrak{R}_0$ is commutative (and so is $\mathfrak{G}$). Since, from Theorem 2, $\mathfrak{R}_0$ is isomorphic to the multiplicative group of the planar ternary ring under consideration, we are done (in view of Theorem 4 and the first part of this theorem).

From Theorem 10 and the example of type (3) given at the beginning of this section, we have a necessary and sufficient condition that π possess a linear coordinate ring with the left distributive law and associative multiplication: π must be of type (3) with two of the $\mathfrak{R}_i$ normal in $\mathfrak{G}$. Following the remark at the end of the example, this is the same as the condition that it possess a linear ring with associative multiplication and the right distributive law (note however, that these are different rings: the points $(0,0)$ and (∞) are interchanged). In fact, if (R, F) is a planar ternary ring which is linear, left distributive, and has associative multiplication, then the interchange of the roles of the points $Y = (\infty)$ and $O = (0,0)$ gives rise to a new ring with the same properties, except that "right distributive"

replaces "left distributive." Thus suppose π is a plane coordinatized by a left near-field (which is not a field); performing the above interchange we arrive at a new system, which however cannot have associative addition. For if it did it would be a right near-field, and it is elementary to see that there would be a collineation moving the line at infinity (L_∞) of the original left near-field coordinate system. It is well-known (see, for instance, [8]) that this implies that the original left near-field is a field, which is contradictory.

The above observations are of interest mainly because they indicate the existence of finite non-trivial examples of what might be called "right (or left) planar division neo-rings": linear planar ternary rings with the right (or left) distributive law, whose addition is not necessarily associative. (In this connection, see [6, 11].) Although the examples given here do not lead to new projective planes, others might, and the lack of associative addition implies that we are not immediately restricted to cases where the order is a prime-power.

We return to the general case of the section now. If ϕ is an automorphism of the group $\mathfrak{G}$ and if $D\phi = Db$ for some $b \in \mathfrak{G}$, then, following Hall ([9]) we call ϕ a *multiplier* of the partial difference set (according to [5], ϕ would be a right multiplier). The concept of multipliers has been a powerful one in the treatment of cyclic difference sets (i. e., cyclic groups of type (0)) and has already been applied to partial difference sets by Hoffman ([10]) and the author ([11]) for types (1b) and (3), respectively; using the techniques of [11], Hoffman's results can be extended from the cyclic case to the abelian case, by the way. Specifically, all of these results are of the following nature: if $\mathfrak{G}$ is abelian, of type (1b) or (3) (or even of type (0)), and if p is any prime divisor of n, then the mapping $\phi\colon x \to x^p$ is a multiplier. Using this result, abelian examples of type (1b) or (3) appear very likely to be of prime-power order (i. e., n is a power of a prime). However, the same mapping is not even one-to-one for abelian examples of type (1a) or (2), and other difficulties present themselves: in the known proofs for the existence of multipliers, a key step is to show that the element $\Delta = \sum d^{-1}$, for all $d \in D$, is a non-singular element of the group algebra of $\mathfrak{G}$ over the rationals. For types (1a), and (2), Δ is definitely singular, and it does not seem unlikely that this should indicate some kind of fundamentally different situation.

Now we investigate the conditions under which we can construct the plane from the group $\mathfrak{G}$. In each of the types all of the left cosets $d\mathfrak{R}_i$, $d \in D$, are distinct. For if $d_1 = d_2 r$, $r \in \mathfrak{R}_i$, $d_1, d_2 \in D$, then $r = d_2^{-1} d_1$ and by Theorem 5, this implies $d_1 = d_2$. Thus, referring to the table at the end of

Section 3, we see that every left coset of $\mathfrak{R}_i$ is of the form $d\mathfrak{R}_i$, $d \in D$, for $i = 0$ in types (1a) and (2), and that all but one left coset of $\mathfrak{R}_i$ is of this form for the remaining values of i. In each of these latter cases, let $q_i \in \mathfrak{G}$ be so chosen that $q_i\mathfrak{R}_i$ is the unique coset not of the form $d\mathfrak{R}_i$; although q_i is not unique, it is determined up to a right multiple by an element of $\mathfrak{R}_i$.

THEOREM 11. *Whenever q_i is defined, $q_i\mathfrak{R}_i = \mathfrak{L}_i q_i$.*

Proof. Choose $b \in \mathfrak{G}$ such that R_i is on $L_i b$. Then $b^{-1}\mathfrak{L}_i b = \mathfrak{R}_i$, or $\mathfrak{L}_i b = b\mathfrak{R}_i$. But $R_i b^{-1}$ is on L_i, so the collection of lines $J_0\mathfrak{R}_i b^{-1}$ all pass through $R_i b^{-1}$; they are all ordinary lines, so none of them contains Q_i, and hence none of them contains P_0. Thus $1 \notin D\mathfrak{R}_i b^{-1}$, or $b \notin D\mathfrak{R}_i$; so $b\mathfrak{R}_i$ is the (unique) left coset of $\mathfrak{R}_i$ which is not of the form $d\mathfrak{R}_i$, $d \in D$. Thus $b\mathfrak{R}_i = q_i\mathfrak{R}_i$, and $q_i = br$, $r \in \mathfrak{R}_i$. So $\mathfrak{L}_i q_i = \mathfrak{L}_i br = b\mathfrak{R}_i r = b\mathfrak{R}_i = q_i\mathfrak{R}_i$.

Indeed, Theorem 11 can be extended to show that $\mathfrak{L}_i q_i$ is the unique right coset of $\mathfrak{L}_i$ which is not representable as $\mathfrak{L}_i d$, $d \in D$. However, we do not need this in what follows.

THEOREM 12. *Besides* (a), (b), (c) *of Theorem* 5, *the following is also satisfied:*

(d) $q_i\mathfrak{R}_i = \mathfrak{L}_i q_i$, where *$q_i\mathfrak{R}_i$ is defined as the unique left coset of $\mathfrak{R}_i$ not representable as $d\mathfrak{R}_i$, $d \in D$, if such a one exists; furthermore, all of the left cosets $d\mathfrak{R}_i$, $d \in D$, are distinct.*

Now if $\mathfrak{G}$ is a group with a subet D and subgroups $\mathfrak{R}_i$ and $\mathfrak{L}_i$, satisfying the numerical conditions of the table at the end of Section 3 for one of the types (1a), (1b), (2), or (3), and if furthermore (a), (b), (c), (d) of Theorems 5 and 12 are satisfied, then we shall say that $(\mathfrak{G}, \mathfrak{R}_i, \mathfrak{L}_i, D)$ is a *partial difference system* (of the appropriate type). We will now show that the existence of a partial difference system implies the existence of a projective plane of the appropriate type. Actually, we will define the plane for each type but only demonstrate that it is a projective plane for type (2), since this is fairly typical of each of the types; the demonstration for the other types is straightforward.

We define a set π of points and lines, with an incidence relation as below.

Points: (a), for each $a \in \mathfrak{G}$; $(\mathfrak{R}_i a)$ for each right coset of $\mathfrak{R}_i$; Q_i, where i runs over the appropriate integers.

Lines: $[Db]$, for each $b \in \mathfrak{G}$; $[\mathfrak{L}_j b]$, for each right coset of $\mathfrak{L}_i$; K_j, where j runs over the appropriate integers.

Incidence:

(a) on $[Db]$ if $a \in Db$; (a) on $[\mathfrak{L}_j b]$ if $a \in \mathfrak{L}_j b$; (a) never on K_j.

$(\mathfrak{R}_i a)$ on $[Db]$ if $b \in \mathfrak{R}_i a$; $(\mathfrak{R}_i a)$ on K_j if $i = j$; $(\mathfrak{R}_i a)$ on $[\mathfrak{L}_j b]$ as below:

Type (1a): never.

Type (1b): if $q_0 a \in \mathfrak{L}_0 b$.

Type (2): never, if $i = 0$ or $j = 0$; for $i = j = 1$, then if $q_1 a \in \mathfrak{L}_1 b$.

Type (3): never, if $i \neq j$; for $i = j$, then if $q_i a \in \mathfrak{L}_i b$.

Q_i never on $[Db]$; Q_i on $[\mathfrak{L}_j b]$ if $i = j$; Q_i on K_j as demanded by π_0.

Now we consider type (2) and show that π, as defined above, forms a projective plane.

Let (a) and (b) be distinct points (i. e., $a \neq b$). Then either $ab^{-1} \in \mathfrak{L}_j$ for some j, or $ab^{-1} = d_1 d_2^{-1}$ for a unique pair $d_1, d_2 \in D$. In the first case we have $a, b \in \mathfrak{L}_j b$, so both (a) and (b) are on $[\mathfrak{L}_j b]$. In the second case we have $d_1^{-1} a = d_2^{-1} b = c$, whence $a, b \in Dc$ and both points are on $[Dc]$. The arguments are easily reversed to show uniqueness.

Consider the points (a) and $(\mathfrak{R}_i b)$. If $i = 0$, then ab^{-1} is in a coset $d\mathfrak{R}_0$, where $d \in D$, so $ab^{-1} = dr$, where $r \in \mathfrak{R}_0$. Let $c = rb$; then $a \in Dc$ and $c \in \mathfrak{R}_0$, so (a) and $(\mathfrak{R}_0 b)$ are on $[Dc]$. If $i = 1$ then either $ab^{-1} \in d\mathfrak{R}_1$ for some $d \in D$, or $ab^{-1} \in q_1 \mathfrak{R}_1$. The first case is handled exactly the same as when $i = 0$. In the second case we have $ab^{-1} \in q_1 \mathfrak{R}_1 = \mathfrak{L}_1 q_1$, and we let $c = q_1 b$. Then $a \in \mathfrak{L}_1 c$ and $q_1 b \in \mathfrak{L}_1 c$, so both points are on the line $[\mathfrak{L}_1 c]$. Again the arguments are readily reversed to demonstrate uniqueness.

Clearly the points (a) and Q_i are on $[\mathfrak{L}_i a]$, and on no other.

Let $(\mathfrak{R}_i a)$ and $(\mathfrak{R}_j b)$ be distinct points (i. e., $\mathfrak{R}_i a \neq \mathfrak{R}_j b$). Then if $i = j$ both points are on the line K_i. Let $i \neq j$, and note that $\mathfrak{R}_j \mathfrak{R}_i = \mathfrak{G}$, since the two subgroups intersect in the identity; for the same reason, every element of $\mathfrak{G}$ has a unique representation in the form $r_j^{-1} r_i$, where $r_j \in \mathfrak{R}_j$, $r_i \in \mathfrak{R}_i$. Let $ba^{-1} = r_j^{-1} r_i$; then $r_i a = r_j b = c$, whence $c \in \mathfrak{R}_i a$, $c \in \mathfrak{R}_j b$, and so $(\mathfrak{R}_i a)$ and $(\mathfrak{R}_j b)$ are both on $[Dc]$. Again, the uniqueness is straightforward.

Consider the points $(\mathfrak{R}_i a)$ and Q_j. If $i \neq j$, then both points are on K_i, while if $i = j = 0$, both points are on K_0. If $i = j = 1$, then let $b = q_1 a$; we have $q_1 a \in \mathfrak{L}_1 b$, so $(\mathfrak{R}_1 a)$ is on $[\mathfrak{L}_1 b]$ and certainly Q_1 is on $[\mathfrak{L}_1 b]$. Uniqueness is obvious.

Finally, the points Q_0 and Q_1 are both on K_0.

Since π is finite and every line clearly contains $n + 1$ points, this is sufficient to prove that π is a projective plane (the nondegeneracy is trivial if $n > 1$). In order to show that π is of type (2), consider mappings of the

form $(a) \to (ax)$, $(\mathfrak{R}_i a) \to (\mathfrak{R}_i ax)$, etc., for each $x \in \mathfrak{G}$. The set of such mappings forms a group of collineations (isomorphic to $\mathfrak{G}$) with respect to which π is of type (2).

Another kind of example of a partial difference system can be constructed as follows. Let R be a finite field of order $n = 2^t$, and let p be any integer such that both p and $p - 1$ are prime to $n - 1$ (e.g., $p = 2$). Let $\mathfrak{G}$ be the group whose elements are the ordered pairs (a, b), $a, b \in R$, $a \neq 0$, with composition $(a, b)(c, d) = (ac, bc + d)$ (then in fact, $\mathfrak{G}$ is isomorphic to the holomorph of the additive group of R with the automorphism group of multiplications). The identity of G is $(1, 0)$, and $(a, b)^{-1} = (a^{-1}, ba^{-1})$. Let D be the subset of $\mathfrak{G}$ consisting of all elements (x, x^p), $x \neq 0$. Then, as both p-th and $(p-1)$-th roots exist and are unique in R, it is easy to see that every element of $\mathfrak{G}$, excepting the elements in the subgroups $\{(1, b)\} = \mathfrak{R}_0 = \mathfrak{L}_0$, and $\{(a, 0)\} = \mathfrak{R}_1 = \mathfrak{L}_1$, can be represented uniquely as $\delta_1\delta_2^{-1}$ and as $\delta_3^{-1}\delta_4$, where the δ_i are in D. Further, (c) and (d) of Theorems 5 and 12 are trivially satisfied.

Thus we have a plane π of type (2), of order $n = 2^t$, with nonabelian $\mathfrak{G}$. Since $\mathfrak{R}_0$ is normal, π possesses a coordinate ring which is linear and has associative addition. However, for $n = 2, 4, 8$, π must be Desarguesian (since all planes of order $\leqq 8$ are known to be Desarguesian), although the author does not know if this is generally true. Also, this gives an example of type (2) in which not all of the subgroups $\mathfrak{R}_i$ are normal (since $\mathfrak{R}_1$ is certainly not).

5. Type $(4, m)$. Throughout this section we asume that π is of type $(4, m)$, where $m \geqq 3$. Since J_0 can be chosen so that R_0, P_0, Q_0 are collinear, we can assume that $\mathfrak{R}_0 = \mathfrak{L}_0$ and (unless stated to the contrary) we shall always make this assumption. From Lemma 2, all the $\mathfrak{R}_i$ and $\mathfrak{L}_i$, for $i \neq 0$, are conjugates, and hence $\mathfrak{G}$ is certainly not abelian, nor are any of the $\mathfrak{R}_i$ or $\mathfrak{L}_i$, $i \neq 0$, even normal in $\mathfrak{G}$: for $\mathfrak{R}_i \cap \mathfrak{R}_j = \mathfrak{L}_i \cap \mathfrak{L}_j = 1$ if $i \neq j$, and $\mathfrak{R}_i$ or $\mathfrak{L}_j$ do not have order one (see the table in Section 3). Let $q = (n-1)/(m-2)$; we shall show that q is an integer.

THEOREM 13. *If $i \neq 0$ and $a \in \mathfrak{G}$, then $a^{-1}\mathfrak{R}_i a$ fixes a subplane π_1 of π, of order $m-1$; $a^{-1}\mathfrak{R}_i a$ fixes exactly $m-2$ tangent points on any line K_j, $j \neq 0$, and q is an integer. $\mathfrak{R}_i$ has exactly q distinct conjugates, any two of which intersect in the identity, and any two of which fix different subplanes of order $m-1$.*

Proof. Let π_1 be the set of points and lines that are fixed by every

element of $a^{-1}\mathfrak{R}_i a$. π_1 contains at least one tangent point on K_i, and so if $j \neq 0$, $a^{-1}\mathfrak{R}_i a$ fixes at least one tangent point on K_j, because there is at least one point Q_k on K_0 which is not on K_i or K_j. Thus π_1 contains at least four points, no three of them collinear, since π_1 certainly contains the m points Q_j, $j \neq 0$; so π_1 is a non-degenerate subplane of π. But π_1 contains only the m points Q_j, $j \neq 0$, from among the points on K_0, and so π_1 must contain m points on all of its lines: thus its order is $m-1$.[3] Besides Q_0 and Q_j, $j \neq 0$, π_1 contains then exactly $m-2$ tangent points on the line K_j. If $j = i$, then these $m-2$ points are just the points $R_i ax$, where x is in the normalizer of $a^{-1}\mathfrak{R}_i a$ in $\mathfrak{G}$. If $a^{-1}\mathfrak{R}_i a \neq b^{-1}\mathfrak{R}_i b$, but there is a non-identity element in common to these conjugate subgroups, then this element must fix exactly m points on K_0, but more than m points on K_i, since it fixes the points that are fixed by $a^{-1}\mathfrak{R}_i a$ and also the points fixed by $b^{-1}\mathfrak{R}_i b$, and these two sets of points (on K_i) cannot be the same (for otherwise the two subgroups would be equal).

Thus distinct conjugates of $\mathfrak{R}_i$ intersect in the identity, and in fact, fix sets of tangent points on K_i which are distinct. So the $n-1$ tangent points on K_i break up into sets of $m-2$ points each, hence q is an integer. In fact, since there are q sets of such tangent points on K_i and each corresponds to a different conjugate of $\mathfrak{R}_i$, q must be the index of the normalizer of $\mathfrak{R}_i$ in $\mathfrak{G}$.

If $\mathfrak{A}$ is any subgroup of $\mathfrak{G}$, let $N(\mathfrak{A})$ be the normalizer of $\mathfrak{A}$ in $\mathfrak{G}$. Using the new parameter q in place of n, $\mathfrak{G}$ has order $q(q-1)(m-2)^2$, $\mathfrak{R}_0$ (and $\mathfrak{L}_0$) has order $q(m-2)$, $\mathfrak{R}_i$ and $\mathfrak{L}_i$, $i \neq 0$, have order $(q-1)(m-2)$, $N(\mathfrak{R}_i)$ and $N(\mathfrak{L}_i)$ have order $(q-1)(m-2)^2$. Since π, of order n, possesses a subplane of order $m-1$, we must have $n = (m-1)^2$ or $n \geqq (m-1)^2 + (m-1) = m^2 - m$. In the first case we have $q = m$ and in the second case $q \geqq (m^2 - m - 1)/(m-2) = m + 1 + 1/(m-2)$; so if $q \neq m$, then $q \geqq m + 2$.

THEOREM 14. *Both m and n are odd, unless $n = (m-1)^2$.*

Proof. If m is even, then $n - 1 \equiv m - 2 \equiv 0 \pmod 2$, so n is odd and $n - m + 1$ is even. So $\mathfrak{R}_i$, $i \neq 0$, possesses an element of order two, and so does each of its conjugates. Thus by Lemma 3, every conjugate of $\mathfrak{R}_i$ must be an $\mathfrak{R}_j$, where $j \neq 0$, and so $q = m$ and $n = (m-1)^2$.

If m is odd, then $n - m + 1 \equiv n \pmod 2$ so either $\mathfrak{R}_0$ or $\mathfrak{R}_i$, $i \neq 0$,

[3] For if $a^{-1}\mathfrak{R}_i a$ fixed a point S on K_0, $S \neq Q_i$ for any i, then since $a^{-1}\mathfrak{R}_i a$ fixes a tangent point T on K_i, $a^{-1}\mathfrak{R}_i a$ would fix the line $L = ST$; but L is an ordinary line, and this is contradictory.

has even order and possesses an element of order two. If n is even then it must be $\mathfrak{R}_i$, $i \neq 0$, that possesses this element, so as above, $q = m$ and $n = (m-1)^2$.

Now let $\mathfrak{H} = N(\mathfrak{R}_1) \cap \mathfrak{R}_0$.

THEOREM 15. *The order of $\mathfrak{H}$ is $m-2$.*

Proof. Consider the $m-2$ points R_1a, where $a \in N(\mathfrak{R}_1)$. Let $i \neq 0, 1$, and let S be the intersection of the lines L_0 and R_1Q_i. Each of the points $L_0(R_1a \cdot Q_i)$, for $a \in N(\mathfrak{R}_1)$, is a point Sb, where $b \in \mathfrak{R}_0$; but clearly, for each such b, we also have $b \in N(\mathfrak{R}_1)$, so $b \in \mathfrak{H}$. Conversely, any element in $\mathfrak{H}$ must be one of the elements b defined in this way, so $\mathfrak{H}$ has order $m-2$.

THEOREM 16. *If $n \neq (m-1)^2$, then both $\mathfrak{H}$ and $\mathfrak{R}_0$ are normal in $\mathfrak{G}$. Furthermore, every element of $\mathfrak{R}_0$ which is not in $\mathfrak{H}$ has order two and $q = 2^t$ for some t.*

Proof. Suppose $b \in \mathfrak{R}_0$, $b \notin \mathfrak{H}$, and $r \in \mathfrak{R}_1$, $r \neq 1$; suppose also that $b^{-1}rb = r$. Then r fixes the point R_1b, whereas (since $b \notin N(\mathfrak{R}_1)$) R_1b is not one of the $m-2$ tangent points on K_1 that are fixed by $\mathfrak{R}_1$. This is impossible (see the proof of Theorem 13). So if $x, y \in \mathfrak{R}_1$ and $x^{-1}bx = y^{-1}by$, we have $b^{-1}(yx^{-1})b = yx^{-1}$, and thus $x = y$. Hence all the $(q-1)(m-2)$ conjugates of b by elements of $\mathfrak{R}_1$ are distinct. Now if $n \neq (m-1)^2$ then $\mathfrak{R}_0$ has even order $n-1$ while $\mathfrak{H}$ has odd order $m-2$, by Theorem 14; so $\mathfrak{R}_0$ possesses an element b of order two, and $b \notin \mathfrak{H}$. The $(q-1)(m-2)$ conjugates of b by elements of $\mathfrak{R}_1$ are all distinct and they all have order two, so by Lemma 3 they must all be in $\mathfrak{R}_0$ (for $\mathfrak{R}_i$, $i \neq 0$, has odd order). So all the elements of $\mathfrak{R}_0$ not in $\mathfrak{H}$ have order two, and there are no other elements of order two in $\mathfrak{G}$. Then any two conjugates of $\mathfrak{R}_0$ must intersect in a group containing these $(q-1)(m-2)$ elements of order two; it is easy to see that this set of elements generate $\mathfrak{R}_0$, so $\mathfrak{R}_0$ is normal in $\mathfrak{G}$. The group $\mathfrak{H}$ consists exactly of all the elements of $\mathfrak{R}_0$ which have odd order (plus the identity), so $\mathfrak{H}$ is normal in $\mathfrak{R}_0$, and even in $\mathfrak{G}$, since $\mathfrak{H}$ is in fact an invariant subgroup of $\mathfrak{R}_0$. Finally, $\mathfrak{R}_0/\mathfrak{H}$ has order q and must also be elementary abelian of exponent two; this finishes the proof.

Now we prove a lemma about arbitrary groups which enables us to classify further our partial difference system.

LEMMA 4. *If G is a (finite or infinite) group, and if H is the subgroup generated by all of the elements which do not have order two, then either $H = 1$, $H = G$, or H has index two in G.*

Proof. Suppose $H \neq 1$, $H \neq G$. If $a \in G$, $a \notin H$, then a has order two; if furthermore, $b \in G$, $b \notin H$, then if $ab \notin H$, $abab = 1$, so $ab = b^{-1}a^{-1} = ba$. If $h \in H$, then since $ah \notin H$, $(ah)^2 = 1$, so $aha = h^{-1}$. Now suppose $a, b \in G$, $a, b, ab \notin H$. For any $h \in H$,

$$h^{-1} = (ab)h(ab) = (ba)h(ab) = b(aha)b = bh^{-1}b = h;$$

i.e., for any $h \in H$, $h^2 = 1$. Since H certainly contains elements (different from the identity) whose order is not two, this is contradictory, whence if $a, b \in G$, $a, b \notin H$, then $ab \in H$. Thus H has index two in G.

THEOREM 17. *In type* $(4, m)$, *if* $m \neq 3$, *then* $n = (m-1)^2$.

Proof. From Theorem 16, $\mathfrak{H}$ consists of all the elements of $\mathfrak{R}_0$ whose order is not two (plus the identity), if $n \neq (m-1)^2$. So, by Lemma 4 either $\mathfrak{H} = 1$, which means $m = 3$, or $\mathfrak{H}$ has index two in $\mathfrak{R}_0$, since $\mathfrak{H}$ is certainly not equal to $\mathfrak{R}_0$. So if $m \neq 3$, then q, which is the index of $\mathfrak{H}$ in $\mathfrak{R}_0$, is two, and this is impossible, since $q \geqq m$.

Thus we have only two cases left. If $n \neq (m-1)^2$, then $m = 3$ and $n = 2^t + 1$, and a good deal more can be said about the group $\mathfrak{G}$ and the plane π. For $\mathfrak{R}_0$ is an elementary abelian group of order 2^t and is normal in $\mathfrak{G}$; for $i \neq 0$, $\mathfrak{R}_i$ defines a transitive and regular automorphism group of $\mathfrak{R}_0$. Thus $\mathfrak{R}_0$ is isomorphic to the additive group of a right near-field (which might be a field), $\mathfrak{R}_1$, say, is isomorphic to the multiplicative group of the same near-field, and $\mathfrak{G}$ is isomorhpic to the holomorph of the additive group of the near-field with the automorphism group of right multiplications. Furthermore, since π contains subplanes of order $m - 1 = 2$, and since π has odd order, π is never Desarguesian. We shall return to this case later to give a more complete description of $\mathfrak{G}$ as the holomorph mentioned above.[4]

Now we shall investigate the converse problem; i.e., find the "axioms" corresponding to Theorem 12, for type $(4, m)$. It is apparent that part of (d) still holds: all of the cosets $d\mathfrak{R}_i$, $d \in D$, are distinct. Since we have chosen R_0 on L_0, it is easy to prove that the unique left coset of $\mathfrak{R}_0$ which is not of the form $d\mathfrak{R}_0$, $d \in D$, is $\mathfrak{R}_0$ itself. In what follows, it is assumed

[4] Let H be a finite group, written additively, of order k, and H' a group of automorphisms of H, transitive and regular on the non-zero elements of H (whence H' has order $k-1$). Let $e \neq 0$ be any fixed element of H, and define a multiplication in H as follows: (i) $x0 = 0$, all $x \in H$; (ii) if $b \neq 0$, then $xb = x\phi_b$, all $x \in H$, where ϕ_b is the unique element of H' satisfying $e\phi_b = b$. Then it is easy to prove that H, under these two operations, is a right near-field. If G is a group of order $k(k-1)$ containing H and H', such that each automorphism of H by an element of H' is an inner automorphism in G, then G is necessarily the holomorph of H with H'.

that i, j, k are all non-zero. If $i \neq j$, choose $g_{ij} \in \mathfrak{G}$ such that the point R_j is on $L_i g_{ij}$; then $\mathfrak{R}_j = g_{ij}^{-1} \mathfrak{L}_i g_{ij}$. Then since P_0 is on L_j, $P_0 g_{ji}$ is on $L_j g_{ji}$, and since R_i is also on $L_j g_{ji}$, $P_0 g_{ji}$ is not on J_0; for if it were, it would have to be the point R_i. Thus $P_0 g_{ji} r$, for any $r \in \mathfrak{R}_i$, is on $L_j g_{ji} r = L_j g_{ji}$ and is not on J_0. So $P_0 g_{ji} r \neq P_0 d$ for any $d \in D$ and any $r \in \mathfrak{R}_i$, and hence $g_{ji} \mathfrak{R}_i \neq d \mathfrak{R}_i$ for any $d \in D$. By a similar argument, it is easy to demonstrate that all of the left cosets $g_{ji} \mathfrak{R}_i$ (as j varies) are distinct. Then, by counting, it is evident that each left coset of $\mathfrak{R}_i$ which is not of the form $d\mathfrak{R}_i$, $d \in D$, is of the form $g_{ji}\mathfrak{R}_i$ for a unique j. Thus we have the result corresponding to (d) of Theorem 12. But we need more here.

Consider a pair of points $R_i b$ and R_j, $i \neq j$, and let L be the line joining them. Then either L is a line $L_k g_{kj}$ for a unique k (i.e., Q_k, $R_i b$, R_j are collinear), or L is a line $J_0 x$. In the first case, R_i is on $L_k g_{ki}$, so $R_i b$ is on $L_k g_{ki} b = L_k g_{kj}$; hence $g_{kj} \in \mathfrak{L}_k g_{ki} b$, or $g_{kj} \mathfrak{R}_j = g_{ki} \mathfrak{R}_i b$, where k is unique. In the second case, since $R_i b$ is on $J_0 x$, we have $R_i b = R_i x$, and similarly, $R_j x = R_j$. So $x \in \mathfrak{R}_i b \cap \mathfrak{R}_j$, and clearly x is the only element in this intersection.

THEOREM 18. *Besides* (a), (b), (c) *of Theorem* 5, *the following are satisfied:*

(e) *There are elements $g_{ij} \in \mathfrak{G}$ for $i, j \neq 0$, $i \neq j$, such that every left coset of $\mathfrak{R}_i$, $i \neq 0$, can be represented uniquely as $d\mathfrak{R}_i$, $d \in D$, or as $g_{ji}\mathfrak{R}_i$, and also having the property $g_{ji}\mathfrak{R}_i = \mathfrak{L}_j g_{ji}$. Every left coset of $\mathfrak{R}_0$ can be expressed uniquely as $d\mathfrak{R}$, $d \in D$, excepting the coset $\mathfrak{R}_0$ itself.*

(f) *If $a \in G$, then $\mathfrak{R}_i a \cap \mathfrak{R}_0$, $i \neq 0$, contains a single element. If $i, j \neq 0$, $i \neq j$, then either $\mathfrak{R}_i a \cap \mathfrak{R}_j$ contains a single element or $g_{ki}\mathfrak{R}_i a = g_{kj}\mathfrak{R}_j$ for a unique k, but not both.*

Proof. Everything is proven, excepting the first sentence of (f). But this is trivial, since different element of $\mathfrak{R}_0$ cannot be in the same left coset of $\mathfrak{R}_i$.

We are now in a position to define the projective plane from a partial difference system of type $(4, m)$ (i.e., a system $(\mathfrak{G}, \mathfrak{R}_i, \mathfrak{L}_i, D)$ satisfying the numerical conditions of the table at the end of Section 3 and the conditions (a), (b), (c), (e), (f) of Theorems 5 and 18). It is worth noting that all of these conditions are not independent; certainly (b) is not needed, for instance. The plane is defined to consist of points and lines exactly as in Section 4, and incidence is also exactly as in Section 4, excepting for the case $(\mathfrak{R}_i a)$ on $[\mathfrak{L}_j b]$, which is as follows:

(i) if $i = j = 0$, then if $\mathfrak{R}_0 a = \mathfrak{L}_0 b$; (ii) never, if one of i or j is zero and the other is not zero; (iii) if $i, j \neq 0$, then if $i \neq j$ and $g_{ji} a \in \mathfrak{L}_j b$.

We shall not carry out the proof that the set of points and lines defined in this fashion, with this incidence relation, forms a projective plane of type $(4, m)$, since it is very similar to the proof given for type (2) in Section 4.

At this point two examples of planes of type $(4, m)$ will be given, both of which have the property $n = (m-1)^2$.

Example 1. Let $m = 3$, $n = 4$. Then π is Desarguesian, and any planar ternary ring for π is isomorphic to $GF(4)$. For each $a \in GF(4)$, $a \neq 0$, define the mapping ϕ_a as follows:

$$\phi_a\colon\ (x, y) \to (ax, ay), \qquad \phi_a\colon\ [m, k] \to [m, ak],$$
$$\phi_a\colon\ [\infty, (k, 0)] \to [\infty, (ak, 0)],$$

where ϕ_a fixes the remaining elements of π. The set $\mathfrak{G}_1$ of such mappings is a collineation group of order three. Furthermore, define the mapping θ as follows:

$$\begin{aligned} \theta\colon\ (x, y) &\to (x^2, y^2) \qquad & [m, k] &\to [m^2, k^2] \\ (m) &\to (m^2) & [\infty, (k, 0)] &\to [\infty, (k^2, 0)], \end{aligned}$$

where again the remaining elements are fixed. Then θ is also a collineation and it is easy to see that group $\mathfrak{G}$ generated by $\mathfrak{G}_1$ and θ is non-abelian of order 6. Furthermore, each element of $\mathfrak{G}$ fixes the points $(0, 0)$, (0), (1), (∞), and the lines joining these points: this is π_0. It is not hard to show that $\mathfrak{G}$ is transitive and regular on both ordinary points and ordinary lines. The subgroup $\mathfrak{R}_0$ is $\mathfrak{G}_1$, and the various $\mathfrak{R}_i$ and $\mathfrak{L}_i$, $i \neq 0$, are the conjugates of the group of order two generated by θ.

Example 2. Let $m = 4$, $n = 9$. Let the planar ternary ring (R, F) be the left near-field of order 9. Then (see [15]), the center of R is a subfield of order 3, which we will call S; furthermore, the automorphism group G_1 of R is non-abelian, of order 6, and G_1 is transitive and regular on the elements of R which are not in S (and of course G_1 fixes every element of S). For each $\phi \in G_1$, let ϕ also denote the mapping of the plane given below:

$$\begin{aligned} \phi\colon\ (x, y) &\to (x\phi, y\phi) \qquad & [m, k] &\to [m\phi, k\phi] \\ (m) &\to (m\phi) & [\infty, (k, 0)] &\to [\infty, (k\phi, 0)], \end{aligned}$$

where ϕ fixes (∞) and L_∞. Then obviously the set $\mathfrak{G}_1$ of all such mappings ϕ is a collineation group isomorphic to G_1. For each $a \in R$, $a \neq 0$, define the mapping θ_a as follows:

$$\begin{aligned} \theta_a\colon\ (x, y) &\to (ax, ay) \qquad & [m, k] &\to [ama^{-1}, ak] \\ (m) &\to (ama^{-1}) & [\infty, (k, 0)] &\to [\infty, (ak, 0)], \end{aligned}$$

where θ_a fixes (∞) and L_∞. Then the set $\mathfrak{G}_2$ of all such mappings is a collineation group of π. Let $\mathfrak{G}$ be the group of collineations generated by $\mathfrak{G}_1$ and $\mathfrak{G}_2$. Since $\theta_a\phi = \phi\theta_{a\phi}$, and since $\mathfrak{G}_2$ has order 8, $\mathfrak{G}$ has order 48. The set of fixed elements of $\mathfrak{G}$ are the points $(0,0)$, (∞), (s), $s \in S$, and the lines $[\infty, (0,0)]$, L_∞, $[s, 0]$, $s \in S$. The point (x,y) is on a fixed line if and only if $x = 0$ or $sx + y = 0$ for some $s \in S$; i.e., if and only if $x = 0$ or $x^{-1}y \in S$. So the ordinary points are the points (x,y) for which $x \neq 0$ and $x^{-1}y \notin S$. If (x,y), (u,v) are a pair of ordinary points, then $(x,y)\phi\theta_a = (a \cdot x\phi, a \cdot y\phi) = (u,v)$ if and only if $a \cdot x\phi = u$, $a \cdot y\phi = v$. This yields $u^{-1}v = (x^{-1}y)\phi$, and since $u^{-1}v$, $x^{-1}y \notin S$, there is exactly one $\phi \in G_1$ satisfying this equation; then a is uniquely determined from $a = u(x\phi)^{-1} = v(y\phi)^{-1}$. So $\mathfrak{G}$ is transitive and regular on ordinary points, and similarly, is transitive and regular on ordinary lines. Suppose (r) is the point R_0; note that $r \notin S$. Then $\mathfrak{R}_0$ consists of those elements $\theta_a\phi$ for which $(ara^{-1})\phi = r$. If $\mathfrak{R}_0$ is normal in $\mathfrak{G}$ then it is clear that $\mathfrak{R}_0$ fixes every point on K_0 $(= L_\infty)$ and thus we would have $(axa^{-1})\phi = x$, all $x \in R$. But this implies that ϕ is an inner automorphism of the multiplicative group of the near-field, and since the right distributive law is not valid, this implies that $\phi = 1$. Then $axa^{-1} = x$, all $x \in R$, so $a \in S$, and $\mathfrak{R}_0$ has order two; since $\mathfrak{R}_0$ must have order 8, this is a contradiction. Hence we have an example of a partial difference system of type $(4, m)$ in which $\mathfrak{R}_0$ is not a normal subgroup of $\mathfrak{G}$, and so Theorem 16 cannot be extended to the case $n = (m-1)^2$. Interestingly, $\mathfrak{G}$ does possess normal subgroups of order 8: $\mathfrak{G}_2$ is an example (there is nothing contradictory in this of course, for a Sylow 2-group of $\mathfrak{G}$ has order 16).

If instead of a left near-field and the group $\mathfrak{G}_2$, we had used a right near-field and the group $\mathfrak{G}_3 = \{\theta_a\}$, $a \neq 0$, where:

$$\theta_a\text{s}\quad (x,y) \to (xa, ya) \qquad [m,k] \to [m, ka]$$
$$(m) \to (m) \qquad [\infty, (k,0)] \to [\infty, (ka, 0)],$$

where θ_a fixes (∞) and L_∞, then $\mathfrak{R}_0 = \mathfrak{G}_3$ would be normal in $\mathfrak{G}$.

The case where $m = 3$, $n \neq (m-1)^2$ appears more interesting than the case $n = (m-1)^2$, if only because n is not restricted to being a square. Some remarks can be made about the first few possible orders $n = 2^t + 1$: 17 is a prime and no non-Desarguesian planes of prime order are known; 33 is not possible, since it is one of the orders rejected by the Bruck-Ryser result ([7]); 65 is not a prime power, and has the further interesting property that 64 is the smallest number for which there is a near-field of characteristic two which is not a field (see [15]).

As remarked previously, if $n \neq (m-1)^2$, $m = 3$, then $\mathfrak{G}$ must be the

holomorph of the additive group of a right near-field with its group of right multiplications. Now we investigate this in more detail. In what follows, R is a right near-field of order 2^t (not excluding the possibility that R is a field). Then $\mathfrak{G}$ can be represented as the set of all couples (a,b), $a,b\in R$, $a\neq 0$, with the operation $(a,b)(c,d)=(ac,bc+d)$. Then $(1,0)$ is the identity of $\mathfrak{G}$, and $(a,b)^{-1}=(a^{-1},ba^{-1})$. $\mathfrak{R}_0$ is the subgroup consisting of all elements $(1,b)$, and $\mathfrak{R}_1$, say, is the subgroup consisting of all elements $(a,0)$. Then $\mathfrak{L}_i$, $i\neq 0$, is the conjugate of $\mathfrak{R}_1$ consisting of all elements (a,y_ia+y_i), and $\mathfrak{R}_i$, $i\neq 0$, is the conjugate consisting of all elements (a,z_ia+z_i), where the y_i and z_i are fixed elements of R, satisfying $y_i\neq y_j$, $z_i\neq z_j$, if $i\neq j$. Then by the proper choice of the point P_0 and the line J_0 (i.e., R_0 on L_0) we can assume that D consists of all elements (x,xT), $x\neq 0,1$, where T is a one-to-one mapping of the non-zero, non-identity elements of R into the non-zero elements of R. If we choose g_{ij} to be the element $(1,y_i+z_j)$, then our system will be a partial difference system of type $(4,m)$ if we demand the following:

(1) If $a\neq 0,1$, and $b\neq y_ia+y_i$, then $(ax)T+xT=bx$ has a unique solution for x.

(2) If $i\neq j$, then $xT\neq y_jx+z_i$ for any $x\in R$.

(3) For all $x,y\in R$, $xT\cdot y+(xy)T\neq z_iy+z_i$.

From these conditions all of (a), (c), (e), (f) can be proven (we do not really need (b), as pointed out earlier, but it too can be proven from (1), (2), (3)). There is nothing complicated about the proof of these statements, and we omit it.

The author does not know of any examples of mappings T (together with choices of the constants y_i and z_i) which satisfy the above conditions. If R is actually a field then T can be chosen to be a polynomial, which might simplify the search for such a mapping.

One further remark about type $(4,m)$. If $n=(m-1)^2$ and if π is Desarguesian, then a coordinate ring R for π (which must be a field) can be chosen so that one of the subplanes π_1 of Theorem 13 is coordinatized by a subfield S of R, where S has order $m-1$ and R has order $n=(m-1)^2$. Then the points of π_0 can be taken as the points $(0,0)$, (∞), (s), $s\in S$. Every collineation of π is given by a linear transformation, by an automorphism of R, or by a product of these two types (see [14]). Using classical homogeneous coordinates for π it is easy to show that at most $2(n-1)$ collineations of π fix the points of π_0, and so $m=3$ is the only possibility. Perhaps the only point in the demonstration which is not com-

pletely obvious is that there are only two choices for an automorphism which fixes the points of π_0.

6. Type $(5, m)$. Now we assume that π is of type $(5, m)$. As in Section 5, we let $N(\mathfrak{A})$ represent the normalizer in $\mathfrak{G}$ of the subgroup $\mathfrak{A}$ of $\mathfrak{G}$. We note that all of the $\mathfrak{R}_i$ and $\mathfrak{L}_j$, $i, j \neq 0$, are conjugates in $\mathfrak{G}$; finally, let $q = n/m$.

THEOREM 19. *Each $a^{-1}\mathfrak{R}_i a$, $i \neq 0$, fixes a subplane of π, of order m, and q is the index of $N(\mathfrak{R}_i)$ in $\mathfrak{G}$. Any pair of the q distinct conjugates of $\mathfrak{R}_i$ intersects in the identity, and distinct conjugates of $\mathfrak{R}_i$ fix different subplanes of π.*

Proof. Let π_1 be the set of points and lines of π that are fixed by every element of $a^{-1}\mathfrak{R}_i a$, and suppose π_1 contains t points on the line K_j, $j \neq 0$. Then since $a^{-1}\mathfrak{R}_i a$ fixes exactly $m+1$ points on K_0, π_1 has order m (it is evident that π_1 is a non-degenerate subplane; see the proof of Theorem 13), and $t - 1 = m$. Since π_1 contains $t-1$ tangent points on K_j, it contains m tangent points on K_j. Following the proof of Theorem 13, it is easy to see that distinct conjugates of $\mathfrak{R}_i$ fix distinct sets of tangent points on K_j, so q is an integer; similarly, $\mathfrak{R}_i$ has q distinct conjugates, any two of which intersect in the identity, so $N(\mathfrak{R}_i)$ has index q in $\mathfrak{G}$.

LEMMA 5. *If $n \neq m^2$, then m is odd and n is even.*

Proof. If m is even, then since $n = qm$, n is also even, so $\mathfrak{R}_i$, $i \neq 0$, has even order $n - m$. Thus, using Lemma 3 and the same argument as in Theorem 14, each conjugate of $\mathfrak{R}_i$ is an $\mathfrak{R}_j$, $j \neq 0$, so $q = m$ and $n = m^2$. If m and n are both odd, then again $n - m$ is even, and the above argument leads to $n = m^2$.

THEOREM 20. *$\mathfrak{H} = N(\mathfrak{R}_1) \cap \mathfrak{R}_0$ has order m.*

Proof. The proof is almost exactly like that of Theorem 15, and we omit it.

THEOREM 21. *In type $(5, m)$, $n = m^2$.*

Proof. If $n \neq m^2$, then as in Theorem 16, $\mathfrak{H}$ has odd order and $\mathfrak{R}_0$ has even order; all the conjugates of an element of $\mathfrak{R}_0$ which is not in $\mathfrak{H}$, by elements of $\mathfrak{R}_1$, are distinct, and so they all have order two, and are all in $\mathfrak{R}_0$. Hence we prove that $\mathfrak{R}_0$ is normal in $\mathfrak{G}$, and that $\mathfrak{H}$ is the subgroup of $\mathfrak{R}_0$ containing all elements of order not two (plus the identity). So by Lemma 4,

$\mathfrak{H}$ has index two in $\mathfrak{R}_0$, whence $q = 2$. But if $n \neq m^2$, then $n \geqq m^2 + m$ and $q \geqq m + 1 \geqq 3$. Thus we have a contradiction, so $n = m^2$.

Now suppose π is a projective plane coordinatized by the linear planar ternary ring (R, F) and suppose the following hold: (i) R has order $n = m^2$, (ii) R has associative addition, (iii) R contains a subset S of order m and an automorphism group G_1 which fixes each element of S and is transitive and regular on the elements not in S. For each $\phi \in G_1$ define the mapping ϕ of π as follows:

$$\begin{array}{ll} \phi\colon\ (x,y) \to (x\phi, y\phi) & [m,k] \to [m\phi, k\phi] \\ \quad\ (m) \to (m\phi) & [\infty, (k,0)] \to [\infty, (k\phi, 0)] \\ \quad\ (\infty) \to (\infty) & L_\infty \to L_\infty. \end{array}$$

Then the set of all such mappings is a group $\mathfrak{G}_1$ (isomorphic to G_1) of collineations of π, and $\mathfrak{G}_1$ has order $n - m$.

Furthermore, for each $a \in R$, define θ_a as follows:

$$\begin{array}{ll} \theta_a\colon\ (x,y) \to (x, y+a) & [m,k] \to [m, k+a] \\ \quad\ (m) \to (m) & [\infty, (k,0)] \to [\infty, (k,0)] \\ \quad\ (\infty) \to (\infty) & L_\infty \to L_\infty. \end{array}$$

Then the set of all such mappings is a group $\mathfrak{G}_2$ of collineations of π, and $\mathfrak{G}_2$ has order n.

Since $\theta_a\phi = \phi\theta_{a\phi}$, the group $\mathfrak{G}$ generated by $\mathfrak{G}_1$ and $\mathfrak{G}_2$ has order $n(n-m)$. $\mathfrak{G}$ fixes the points (∞), (s), $s \in S$, and the lines L_∞, $[\infty, (s,0)]$, $s \in S$. (This is π_0.) It is easy to see that $\mathfrak{G}$ is transitive and regular on the points of π which are not on lines of π_0, and on the lines of π which do not contain points of π_0. So π is of type $(5, m)$.

In [8] Hall has described a technique for constructing a class of V-W systems, as follows. Let S be a field of order $p^t \neq 2$, p any prime, let $z^2 - az - b$, for $a, b \in S$, be an irreducible quadratic over S, and let R be the set of all elements $\lambda x + y$, for $x, y \in S$, where λ is some indeterminant. Define addition in R by $(\lambda x + y) + (\lambda u + v) = \lambda(x+u) + (y+v)$, and multiplication by:

(i) $y(\lambda u + v) = \lambda yu + yv$,

(ii) if $x \neq 0$, then

$$(\lambda x + y)(\lambda u + v) = \lambda(au + xv - yu) + x^{-1}u(-y^2 + ay + b) + yv.$$

Then R is a left V-W system, does not satisfy the right distributive law, and does not have associative multiplication unless $p^t = 3$, $a = 0$, $b = 2$.

The automorphism group of R consists of all mappings T given by $(\lambda x + y)T = \lambda(x\phi \cdot c) + x\phi \cdot d + y\phi$, where ϕ is any automorphism of S which fixes a, b, and where c, d are arbitrary in S, except $c \neq 0$. (These statements are easy enough to prove given (i) and (ii); (i) and (ii) can be derived from the directions given in [8].)

If we let $\phi = 1$, then we have a group G_1 of automorphisms, of order $p^t(p^t - 1)$, fixing every element of S, transitive and regular on the elements not in S. Thus each V-W system of this class gives an example of a plane of type $(5, m)$. For $p^t = 2$ (i.e., $n = 4$, $m = 2$) it is easy to see that the mapping of $GF(4)$ given by $x \to x^2$ generates a group G_1 with the desired properties. So for every order $n = p^{2t}$, p a prime, there is a plane of type $(5, m)$. Using the argument at the end of Section 5, it can be shown that the only Desarguesian example must have $m = 2$, $n = 4$.

The problem of constructing the plane from the partial difference system is almost exactly like that for type $(4, m)$. The g_{ij} are defined in exactly the same way, whenever Q_i is not on K_j (which is whenever i and j are both non-zero). Since every left coset of $\mathfrak{R}_0$ is of the form $d\mathfrak{R}_0$, $d \in D$, the last sentence of (e) of Theorem 18 is deleted, and the rest remains. (And of course, (a), (b), (c) of Theorem 5 still hold.)

7. Type $(6, m)$. For this type we note that there is nothing "special" about the subgroups $\mathfrak{R}_0$ and $\mathfrak{L}_0$, and thus all of the $\mathfrak{R}_i$ and $\mathfrak{L}_j$ are conjugate in $\mathfrak{G}$. Let $q = (n - m)/(m^2 - m)$.

THEOREM 22. *The order of $N(\mathfrak{R}_i)$ is $(m^2 - m)(n - m^2)$, and q is an integer. Any pair of the q distinct conjugates of $\mathfrak{R}_i$ intersect in the identity. For each conjugate of $\mathfrak{R}_i$ there is a subplane of π, of order m^2, containing π_0, which is fixed (element-wise) by each element of the conjugate; distinct conjugates fix different subplanes.*

Proof. Consider the subgroup $a^{-1}\mathfrak{R}_i a$, and suppose this subgroup fixes t (where t is necessarily positive) tangent points on K_i. Then, using the type of argument found in the proof of Theorem 13, it is clear that $a^{-1}\mathfrak{R}_i a$ fixes a subplane of π, of order $m + t$, and this subplane contains π_0. Also, as in Theorem 13, distinct conjugates fix different subplanes, each of which has the same order $m + t$, and each conjugate fixes a different set of t tangent points on K_i. Since each of these subplanes of order $m + t$ has the property that every one of its points is on a line of π_0, each subplane must have order m^2, so $t = m^2 - m$. Since t must divide $n - m$, which is the number of tangent points on K_i, q must be an integer; furthermore, t is the index

of $\mathfrak{R}_i$ in its normalizer, so the order of $N(\mathfrak{R}_i)$ is $(m^2-m)(n-m^2)$. Again as in Theorem 13, distinct conjugates of $\mathfrak{R}_i$ intersect in the identity.

THEOREM 23. *In type* $(6,m)$, $n=m^4$.

Proof. From Theorem 22, q is an integer, so $n=q(m^2-m)+m\equiv m$ $(\operatorname{mod} 2)$, since m^2-m is even. But then $n-m^2$, which is the order of $\mathfrak{R}_i$, is even, so $\mathfrak{R}_i$ contains an element of order two, and so does each conjugate of $\mathfrak{R}_i$. Then, by Lemma 3, every conjugate of $\mathfrak{R}_i$ is an $\mathfrak{R}_j$, so $q=m^2+m+1$. This immediately implies $n=m^4$.

The author does not know of any example of type $(6,m)$. However, using the argument of the last paragraph of Section 5, it can be shown that no example can be Desarguesian. The problem of defining the plane from the abstractly given partial difference system is also similar to the situation for types $(4,m)$ and $(5,m)$. The g_{ij} are defined in the same way, and then the conditions (besides (a), (b), (c) of Theorem 5) are exactly those of Theorem 18, excepting that the second sentence of (e), the first sentence of (f), and all references that would prevent i or j from being zero, are deleted.

In order to construct a plane of type $(6,m)$, the plane π_0 must be fully known. Since the only planes known at the present time have prime-power order, the use of such a known plane would result in a plane π which also had prime-power order. Thus this type does not seem to offer a very practical method of constructing planes of new orders.

8. Remarks. As pointed out above, type $(6,m)$ appears to be perhaps the least hopeful method of constructing planes of non-prime-power order. For somewhat similar reasons, type $(5,m)$ and type $(4,m)$ with $n=(m-1)^2$ do not appear particularly promising. But type $(4,m)$, with $m=3$, looks quite interesting, and an investigation of the conditions (1), (2), (3) of Section 5 might lead to some new planes. The remaining types seem to call for further study, although new and deeper techniques will probably be necessary.

The existence of right or left planar division neo-rings that are not V-W systems, even though the examples given in this paper define well-known planes, might be investigated further. Without the assumption of associative multiplication, such systems need not lead to planes of type (3). But using the techniques of [11] (see also [6]), it can be shown that if R is a left planar division neo-ring then the right nucleus of the multiplicative loop of R, plus the zero element, forms a subsystem of the same type, with

associative multiplication (the right nucleus of a loop G consists of all $b \in G$ such that $(xy)b = x(yb)$ for all $x, y \in G$). Hence associative systems of this kind form a natural starting point for any investigation.

Finally, it is worth noting that practically every type of finite projective plane known at the present time is included in at least one way in the class of partially transitive planes.

THE OHIO STATE UNIVERSITY.

REFERENCES.

[1] E. Artin, "Coordinates in affine geometry," *Reports of a Mathematical Colloquium*, Issue 2, pp. 15-20, University of Notre Dame, 1945.

[2] R. Baer, "Homogeneity of projective planes," *American Journal of Mathematics*, vol. 64 (1942), pp. 137-152.

[3] ———, "Projectivities of finite projective planes," *ibid.*, vol. 69 (1947), pp. 653-684.

[4] R. C. Bose, "An affine analogue of Singer's Theorem," *Journal of the Indian Mathematical Society*, vol. 6 (1942), pp. 1-15.

[5] R. H. Bruck, "Difference sets in a finite group," *Transactions of the American Mathematical Society*, vol. 78 (1955), pp. 464-481.

[6] ———, "Analogues of the ring of rational integers," *Proceedings of the American Mathematical Society*, vol. 6 (1955), pp. 50-58.

[7] R. H. Bruck and H. J. Ryser, "The non-existence of certain finite projective planes," *Canadian Journal of Mathematics*, vol. 1 (1949), pp. 88-93.

[8] M. Hall, Jr., "Projective planes," *Transactions of the American Mathematical Society*, vol. 54 (1943), pp. 229-277.

[9] ———. "Cyclic projective planes," *Duke Mathematical Journal*, vol. 14 (1947), pp. 1079-1090.

[10] A. J. Hoffman, "Cyclic affine planes," *Canadian Journal of Mathematics*, vol. 4 (1952), pp. 295-301.

[11] D. R. Hughes, "Planar division neo-rings," *Transactions of the American Mathematical Society*, vol. 80 (1955), pp. 502-527.

[12] L. Paige, "Neofields," *Duke Mathematical Journal*, vol. 16 (1949), pp. 39-60.

[13] G. Pickert, *Projektive Ebenen*, Berlin, 1955.

[14] O. Veblen and J. W. Young, *Projective geometry*, New York, 1910.

[15] H. Zassenhaus, "Über endliche Fastkörper," *Abhandlungen aus dem Mathematischen Seminar der Hamburg Universität*, vol. 11 (1935), pp. 187-220.

ON A THEOREM OF LAZARD.*

By JEAN DIEUDONNÉ.

1. M. Lazard has proved, by an ingenious direct argument [3], that any formal Lie group of dimension 1 over a commutative ring K with unit element and without nilpotent elements, is necessarily *abelian*. When K is a field of characteristic 0, Lie theory yields a trivial proof of that result, and it was natural to expect that there should also be a simple proof using Lie theory when K is a field of characteristic $p > 0$. Up to now, however, I had only been able to give (in [2]) such a proof under the additional assumption that $X_0^p \neq 0$. (I use the terminology and notations of [1] and [2]). The purpose of this note is to complete the proof by treating the remaining case $X_0^p = 0$.

2. We recall that the hyperalgebra $\mathfrak{G}$ of the one-dimensional group under consideration has a basis over K consisting of the unit element I and of the monomials $X_\alpha = X_0^{\alpha_0} X_1^{\alpha_1} \cdots X_r^{\alpha_r}$, with $0 \leqq \alpha_i < p$; the elements of that basis other than I generate a two-sided ideal $\mathfrak{G}_+$. Our proof rests on the two following observations:

a) for any pair of elements U, V in $\mathfrak{G}$, $[U, V] \in \mathfrak{G}_+$;

b) a relation of the form $\lambda X_0 + X_0 U = 0$, where $U \in \mathfrak{G}_+$ and $\lambda \in K$, implies $\lambda = 0$: indeed, the product of X_0 with any monomial $X_\alpha \in \mathfrak{G}_+$ is either 0 or a monomial X_β of total degree $\geqq 2$, due to the assumption $X_0^p = 0$.

We have seen in [2, pp. 227-228] that X_0 commutes with every other X_i. Using induction, we assume that $X_0, X_1, \cdots, X_{r-1}$ commute with all the X_i, and that $X_{r+1}, X_{r+2}, \cdots, X_{s-1}$ commute with X_r; all we need to do is to prove that X_s also commutes with X_r.

Using Lemma 1 of [2, p. 224], we have

$$[X_s, X_r] = aX_0 \text{ with } a \in K. \tag{1}$$

From the definition of the Frobenius homomorphism p' [1, p. 103], it

* Received May 28, 1956.

follows that the kernel of $\boldsymbol{p}'$ is generated by all monomials X_α in which $\alpha_0 > 0$; from (1) we derive therefore

$$(2) \qquad [X_{2s}, X_{r+s}] = a^{p^s} X_s + X_{s-1}U_{s-1} + \cdots + X_1 U_1 + X_0 U_0$$

where the U_i belong to $\mathfrak{G}$. Similarly, from the assumption $[X_{i-1}, X_{r-1}] = 0$, it follows that, for every $i \geqq 0$

$$(3) \qquad [X_i, X_r] = X_0 V_i \text{ with } V_i \in \mathfrak{G}.$$

As X_r commutes with $X_0, \cdots, X_{s-1}$, we derive from (2) and (1)

$$(4) \qquad [[X_{2s}, X_{r+s}], X_r]$$
$$= a^{p^s+1} X_0 + X_{s-1}[U_{s-1}, X_r] + \cdots + X_1[U_1, X_r] + X_0[U_0, X_r]$$

and from (3) and the identity $[X, YZ] = [X, Y]Z + Y[X, Z]$ it follows that

$$(5) \qquad [[X_{2s}, X_{r+s}], X_r] = a^{p^s+1} X_0 + X_0 W$$

with $W \in \mathfrak{G}_+$. But, by the Jacobi identity, the left-hand side of (5) is $[X_{2s}, [X_{r+s}, X_r]] - [X_{r+s}, [X_{2s}, X_r]]$; using (3) and remark a), and remembering that X_0 commutes with everything, this expression has also the form $X_0 W'$, with $W' \in \mathfrak{G}_+$. Remark b) then proves that $a = 0$, q. e. d.

NORTHWESTERN UNIVERSITY.

REFERENCES.

[1] J. Dieudonné, "Groupes de Lie et hyperalgèbres de Lie sur un corps de caractéristique $p > 0$," *Comm. Math. Helv.*, vol. 28 (1954), pp. 87-118.

[2] ———, "Lie groups and Lie hyperalgebras over a field of characteristic $p > 0$ (II)," *American Journal of Mathematics*, vol. 77 (1955), pp. 218-244.

[3] M. Lazard, "La non-existence des groupes de Lie formels non abéliens à un paramètre," *C-R. Acad. Sci. Paris*, vol. 239 (1954), pp. 942-945.

A QUALITATIVE CHARACTERIZATION OF BLASCHKE PRODUCTS IN A HALF-PLANE.*[1]

By Edwin J. Akutowicz.

1. Introduction. Let $a_1, a_2, \cdots$ be a sequence of complex numbers with $\operatorname{Im} a_\nu > 0$ for all ν, and such that

$$(1) \qquad \sum_{\nu=1}^{\infty} \operatorname{Im} a_\nu/(1+|a_\nu|^2) < \infty.$$

This condition guarantees that the Blaschke product $b(z)$ with the a_ν as its set of zeros converges uniformly in any bounded closed set contained in the half-plane $\operatorname{Im} z \geqq a$, where a is any positive number, and independently of the order of the factors:

$$(2) \qquad b(z) = ((z-i)/(z+i))^n \prod_{\nu=1}^{\infty} (|a_\nu - i|/(a_\nu - i))(|a_\nu + i|/(a_\nu + i)) \times ((z-a_\nu)/(z-\bar{a}_\nu)),$$

where n is a non-negative integer. *We make the fixed assumption in Section* 1 *and Section* 2 *that* $\operatorname{Im} a_\nu > 0$ *and that* (1) *holds.*

It is familiar that $|b(x+iy)| < 1$ for $y > 0$, and that $|b(x+iy)| \to 1$ for almost all x, $-\infty < x < \infty$, as $y \to 0+$. We shall prove that

$$(3) \qquad \int_{-\infty}^{\infty} \log |b(z)|/(1+x^2)dx \to 0 \text{ as } y \to 0+.$$

(We write $z = x+iy$). Furthermore, we shall prove in Section 3 that any function $F(z)$, holomorphic in the half-plane $y > 0$, satisfying the conditions $|F(z)| < 1$ and

$$\int_{-\infty}^{\infty} \log |F(z)|/(1+x^2)dx \to 0 \text{ as } y \to 0+,$$

is necessarily of the form $F(z) = e^{ikz+ic}b(z)$, where k, c are real constants, $k \geqq 0$, and $b(z)$ is the Blaschke product formed with the zeros of $F(z)$. If one also assumes $(1/r) \log |F(re^{i\theta})| \to 0$ as $r \to \infty$, for some θ, $0 < \theta < \pi$,

* Received April 22, 1955; revised May 14, 1956.

[1] This research was supported jointly by the Army, Navy, and Air Force under contract with the Lincoln Laboratory, Massachusetts Institute of Technology.

then $k=0$. This characterization of $b(z)$ is thus in terms of a combination of general properties which have no explicit reference to the formula (2).

2. Proof of Necessity. The product $b(z)$, given by (2), is convergent for $y>0$ and we can choose a positive real number λ, which will be fixed once and for all, such that $b(i\lambda)\neq 0$. Then $b(iy)$ does not vanish for y in some closed interval $\mathfrak{I}$ centered at λ. We may take $\mathfrak{I}$ such that the length of $\mathfrak{I}$ is smaller than λ. Let ϵ be positive and such that $\lambda-\epsilon$ belongs to $\mathfrak{I}$ (so that $\epsilon\leqq\lambda/2$), and map $\operatorname{Im} z>0$ to $|w|<1$ by $w=(z-i(\lambda-\epsilon))/(z+i(\lambda-\epsilon))$. Then the horizontal line $y=\epsilon$ corresponds to the circle C_ϵ with center $(\epsilon/\lambda,0)$ and radius $(\lambda-\epsilon)/\lambda$. The line element $|dw|$ on C_ϵ is related to the line element dx on $y=\epsilon$ by the equation $|dw|=\{2(\lambda-\epsilon)/(x^2+\lambda^2)\}dx$. Now

$$0<\min(1,\lambda^2)\leqq(\lambda^2+x^2)/(1+x^2)\leqq\max(1,\lambda^2)\quad\text{for }-\infty<x<\infty$$

implies that, as $y\to 0+$,

$$\int_{-\infty}^{\infty}\log|b(z)|/(1+x^2)\,dx\to 0\ \text{ if and only if }\int_{-\infty}^{\infty}\log|b(z)|/(x^2+\lambda^2)\,dx\to 0.$$

Therefore for $\operatorname{Im} z=y=\epsilon$, we have [2] $\displaystyle\int_{-\infty}^{\infty}\log|b(z)|/(x^2+\lambda^2)\,dx$

$$=\int_{-\infty}^{\infty}(x^2+\lambda^2)^{-1}\log\prod_\nu|(z-a_\nu)/(z-\bar a_\nu)|\,dx$$

$$=\sum_\nu\int_{-\infty}^{\infty}(x^2+\lambda^2)^{-1}\log|(z-a_\nu)/(z-\bar a_\nu)|\,dx$$

$$=(2(\lambda-\epsilon))^{-1}\sum_\nu\int_{C_\epsilon}\log|\{[i(\lambda-\epsilon)(1+w)/(1-w)]-a_\nu\}$$
$$/\{[i(\lambda-\epsilon)(1+w)/(1-w)]-\bar a_\nu\}|\,|dw|$$

$$=(2(\lambda-\epsilon))^{-1}\sum_\nu\{(2\pi(\lambda-\epsilon)/\lambda)\log|(i(\lambda-\epsilon)+a_\nu)/(i(\lambda-\epsilon)-a_\nu)|$$
$$+\int_{C_\epsilon}\log|(w-k_\nu)/(w-k'_\nu)|\,|dw|\},$$

where

$$k_\nu=-(i(\lambda-\epsilon)-a_\nu)/(i(\lambda-\epsilon)+a_\nu),$$
$$k'_\nu=-(i(\lambda-\epsilon)-\bar a_\nu)/(i(\lambda-\epsilon)+\bar a_\nu).$$

Then $|k_\nu|<1$ and $|k'_\nu|>1$, and if k_ν lies in the interior or on the boundary of C_ϵ we can use Jensen's formula [5] to obtain

---

[2] We take $n=0$ in (2), which is clearly permissible.

$$\int_{C_\epsilon} \log|(w-k_\nu)/(w-k'_\nu)|\,|dw|$$
$$=((\lambda-\epsilon)/\lambda)\int_{|s|=(\lambda-\epsilon)/\lambda} \log|\{s-(k_\nu-(\epsilon/\lambda))\}/\{s-(k'_\nu-(\epsilon/\lambda))\}|\,d\arg s$$
$$=(2\pi(\lambda-\epsilon)/\lambda)\log|(i(\lambda-\epsilon)-a_\nu)/(i(\lambda+\epsilon)+a_\nu)|.$$

If k_ν lies outside C_ϵ, $\log|(w-k_\nu)/(w-k'_\nu)|$ is harmonic in a neighborhood of the closed disk bounded by C_ϵ and the Gauss mean value theorem yields

$$\int_{C_\epsilon} \log|(w-k_\nu)/(w-k'_\nu)|\,|dw| = 2\pi((\lambda-\epsilon)/\lambda)\log|((\epsilon/\lambda)-k_\nu)/((\epsilon/\lambda)-k'_\nu)|$$
$$=2\pi((\lambda-\epsilon)/\lambda)\log|(i(\lambda+\epsilon)-a_\nu)(i(\lambda-\epsilon)-a_\nu)$$
$$/(i(\lambda+\epsilon)+a_\nu)(i(\lambda-\epsilon)+a_\nu)|.$$

Hence

$$\int_{-\infty}^{\infty}(x^2+\lambda^2)^{-1}\log|b(z)|\,dx=(\pi/\lambda)\sum_\nu\Bigg(\log|(i(\lambda-\epsilon)+a_\nu)/(i(\lambda-\epsilon)-a_\nu)|$$
$$+\begin{cases}\log|(i(\lambda-\epsilon)-a_\nu)/(i(\lambda+\epsilon)+a_\nu)| & \text{if } \operatorname{Im} a_\nu\geqq\epsilon\\ \log|(i(\lambda+\epsilon)-a_\nu)(i(\lambda-\epsilon)-a_\nu)/(i(\lambda+\epsilon)+a_\nu)(i(\lambda-\epsilon)+a_\nu)| & \text{if } \operatorname{Im} a_\nu<\epsilon\end{cases}\Bigg)$$

$$=(\pi/\lambda)\sum_{\operatorname{Im} a_\nu\geqq\epsilon}\log|(i(\lambda-\epsilon)+a_\nu)/(i(\lambda+\epsilon)+a_\nu)|$$
$$+(\pi/\lambda)\sum_{\operatorname{Im} a_\nu<\epsilon}\log|(i(\lambda+\epsilon)-a_\nu)/(i(\lambda+\epsilon)+a_\nu)|. \tag{4}$$

In order to prove that the first sum in (4), which we shall denote by Δ, tends to 0 as $\epsilon\to 0$ we write $b(z)=b_1(z)b_2(z)b_3(z)$, where the zeros of $b_1(z)$ lie in the rectangle $0<y<1$, $|x|<P$, those of $b_2(z)$ in the half-plane $y\geqq 1$, and those of $b_3(z)$ in the two infinite strips $0<y<1$, $|x|\geqq P$. Here P is at our disposal.

The zeros of the product $b_1(z)$ can be enumerated so that

$$y_1\geqq y_2\geqq y_3\geqq\cdots,\ \lim y_\nu=0, \tag{5}$$

where $a_\nu=x_\nu+iy_\nu$, $|x_\nu|<P$, $0<y_\nu<1$.
Then the first sum in (4) for $b=b_1$ contains a finite number of summands, and so is certainly finite. We now prove that it tends to 0 as $\epsilon\to 0+$. Essential use is made of (5). We have

$$\Delta_1=\sum_{y_\nu\geqq\epsilon}\log|(i(\lambda-\epsilon)+a_\nu)/(i(\lambda+\epsilon)+a_\nu)|.$$

Then, for $\epsilon = y_q$, q large,

$$2\Delta_1 = \sum_{\nu=1}^{q} \log\left(\{(\lambda + y_\nu - y_q)^2 + x_\nu^2\}/\{(\lambda + y_\nu + y_q)^2 + x_\nu^2\}\right)$$
$$= \sum_{\nu=1}^{q} \{(4y_q(\lambda + y_\nu)/(\lambda + y_\nu + y_q)^2 + x_\nu^2) \text{ times a uniformly bounded factor}\},$$

by the mean value theorem, plus the fact that $(i(\lambda - \epsilon) + a_\nu)/(i(\lambda + \epsilon) + a_\nu)$ is bounded away from 0, uniformly for all ν and $\epsilon \leqq \lambda/2$. Therefore,

$$|2\Delta_1| \leqq My_q \sum_{\nu=1}^{q} (\lambda^2 + y_\nu^2 + x_\nu^2)^{-1} + My_q \sum_{\nu=1}^{q} y_\nu(\lambda^2 + y_\nu^2 + x_\nu^2)^{-1}.$$

The second sum obviously tends to 0 as $q \to \infty$, and for $q > N$

$$y_q \sum_{\nu=1}^{q} (\lambda^2 + y_\nu^2 + x_\nu^2)^{-1} < y_q \sum_{\nu=1}^{N} (\lambda^2 + y_\nu^2 + x_\nu^2)^{-1} + \sum_{\nu=N+1}^{\infty} y_\nu(\lambda^2 + y_\nu^2 + x_\nu^2)^{-1}.$$

For N sufficiently large the last sum in the preceding line is arbitrarily small, and the rest tends to 0 since $y_q \to 0$ as $q \to \infty$. Therefore $\Delta_1 \to 0$ for any $P > 0$.

We now consider $b_2(z)$, and take $\epsilon < \min(1, \lambda/2)$. For this product

$$0 < -2\Delta_2 = \sum_{y_\nu \geqq 1} \log|(i(\lambda + \epsilon) + a_\nu)/(i(\lambda - \epsilon) + a_\nu)|^2$$
$$= \sum_{y_\nu \geqq 1} \log\{1 + 4\epsilon[(y_\nu + \lambda)/(x_\nu^2 + (y_\nu + \lambda - \epsilon)^2]\}$$
$$< 4\epsilon \sum_{y_\nu \geqq 1} (y_\nu + \lambda)/(x_\nu^2 + (y_\nu + (\lambda/2))^2)$$
$$< 4\epsilon(\lambda + 2) \sum_{y_\nu \geqq 1} y_\nu/(x_\nu^2 + y_\nu^2).$$

Hence $\Delta_2 \to 0$ as $\epsilon \to 0$, since the last sum is finite.

For $b_3(z)$ with $P = 1$ we have

$$\sum_{0 < y_\nu < 1,\ |x_\nu| \geqq 1} y_\nu/(x_\nu^2 + y_\nu^2) < \infty,$$

and hence if $\epsilon' > 0$ we can fix P at a sufficiently large value, *independent of* ϵ, so that

$$\sum_{0 < y_\nu < 1,\ |x_\nu| \geqq P} y_\nu/(x_\nu^2 + y_\nu^2) < \epsilon'.$$

Then, writing $\sum' = \sum_{\epsilon \leqq y_\nu < 1,\ |x_\nu| \geqq P}$, it follows that

$$0 < -2\Delta_3 < 4\epsilon \textstyle\sum' (y_\nu + \lambda)/\{x_\nu^2 + (y_\nu + (\lambda/2))^2\}$$
$$\leqq 4\epsilon(1 + \lambda) \textstyle\sum' 1/(x_\nu^2 + y_\nu^2)$$
$$\leqq 4(1 + \lambda) \textstyle\sum' y_\nu/(x_\nu^2 + y_\nu^2) < 4\epsilon'(1 + \lambda),$$

and this proves that Δ_3 is arbitrarily small. Therefore $\Delta \to 0$ as $\epsilon \to 0$ for any Blaschke product, since $\Delta = \Delta_1 + \Delta_2 + \Delta_3$.

Turning now again to (4), it is easy to see that $\sum_{\operatorname{Im} a_\nu < \epsilon}$ tends to 0 as $\epsilon \to 0+$ for any Blaschke product. (Here we again use the fact that $b(i\lambda) \neq 0$.) Thus (3) holds for any Blaschke product $b(z)$. Hence we have proved

THEOREM 1. *Let $b(z)$ be a Blaschke product in the upper half-plane,* $\operatorname{Im} z = \operatorname{Im}(x + iy) = y > 0$. *Then*

$$\int_{-\infty}^{\infty} (1 + x^2)^{-1} \log |b(z)| \, dx \to 0 \quad as \quad y \to 0+.$$

COROLLARY. *For any finite real numbers l_1 and l_2, $l_1 < l_2$, and any Blaschke product $b(z)$,*

$$\int_{l_1}^{l_2} \log |b(x + iy)| \, dx \to 0 \quad as \quad y \to 0+.$$

This has been established by W. Kryloff [2] by a rather complicated argument. It is a trivial consequence of Theorem 1 and the inequality

$$(1 + \{\max(|l_1|, |l_2|)\}^2) \int_{-\infty}^{\infty} (1 + x^2)^{-1} \log |b(z)| \, dx \leqq \int_{l_1}^{l_2} \log |b(z)| \, dx < 0.$$

3. Proof of sufficiency. In this section we establish the following uniqueness theorem.

THEOREM 2. *Let $F(z)$ be holomorphic for* $\operatorname{Im} z = \operatorname{Im}(x + iy) = y > 0$, *and such that $|F(z)| < 1$ and*

$$\int_{-\infty}^{\infty} (1 + x^2)^{-1} \log |F(z)| \, dx \to 0 \quad as \quad y \to 0+. \tag{6}$$

Then $F(z) = e^{ikz+ic} b(z)$, where k and c are real constants, $k \geqq 0$, and $b(z)$ is the Blaschke product with the zeros of $F(z)$ as its set of zeros.

In the first place we claim that the Blaschke product $b(z)$ formed with the zeros of $F(z)$ is convergent. For we can put $f(w) = F(i(1 + w)/(1 - w))$ to define a bounded holomorphic function $f(w)$ in $|w| < 1$, $w = \rho e^{i\theta}$. The integral of $\log |f(\rho e^{i\theta})|$ over $0 \leqq \theta < 2\pi$ does not decrease when ρ increases, and since $f(w)$ is bounded, it must increase to a finite limit as $\rho \to 1-$. Therefore, by a theorem of A. Ostrowski [3] [3] the Blaschke product [4] in the

[3] This result of Ostrowski is much stronger than what is necessary for our purposes.

[4] This is given by $w^n \prod_\nu (|w_\nu|/w_\nu)((w_\nu - w)/(1 - \bar{w}_\nu w))$, $\{w_\nu\}$ being the set of zeros of f which are different from 0, and n a non-negative integer.

unit circle, formed with the zeros of $f(w)$, is convergent in $|w| < 1$, and this is equivalent to $b(z)$ formed with the zeros of $F(z)$ being convergent in $y > 0$. Let us take the zeros of $F(z)$ in any enumeration: $a_1, a_2, \cdots$. Let $b_m(z)$ denote the m-th partial product of $b(z)$. In modulus, the factors of $b_m(z)$ are $|(z - a_\nu)/(z - \bar{a}_\nu)|$, $\nu = 1, 2, \cdots, m$, and if $\eta > 0$ is given, there exists a positive number $\epsilon_0(m, \eta)$ such that

$$|(x + i\epsilon - a_\nu)/(x + i\epsilon - \bar{a}_\nu)| > 1 - (\eta/m), \text{ for } 0 < \epsilon < \epsilon_0(m, \eta),$$
$$-\infty < x < \infty, \ \nu = 1, 2, \cdots, m.$$

Therefore, $|b_m(x + i\epsilon)| > (1 - (\eta/m))^m > (1 - \eta)$, and

$$|F(x + i\epsilon)/b_m(x + i\epsilon)| < (1 - \eta)^{-1} \text{ for } 0 < \epsilon < \epsilon_0(m, \eta), \ -\infty < x < \infty.$$

By the maximum principle,

$$|F(z)/b_m(z)| < (1 - \eta)^{-1} \text{ for } 0 < \epsilon < \epsilon_0(m, \eta), \ \operatorname{Im} z \geqq \epsilon.$$

This implies, successively, that

$$|F(z)/b_m(z)| < (1 - \eta)^{-1}, \ |F(z)/b(z)| \leqq (1 - \eta)^{-1}, \ |F(z)/b(z)| \leqq 1,$$
$$\text{for } \operatorname{Im} z > 0.$$

Therefore, $\phi(z) \equiv F(z)/b(z)$ is holomorphic, bounded in modulus by 1, and nonvanishing in $y > 0$. By (6) and Theorem 1,

$$(7) \quad \int_{-\infty}^{\infty} (1 + x^2)^{-1} \log |\phi(x + iy)| dx$$
$$= \int_{-\infty}^{\infty} (1 + x^2)^{-1} \log |F(x + iy)| dx - \int_{-\infty}^{\infty} (1 + x^2)^{-1} \log |b(x + iy)| dx \to 0$$

as $y \to 0+$. It is easy to see that $\log |\phi(x + iy)|$ can be written in the form

$$(8) \quad \log |\phi(x + iy)| = -ky - \tfrac{1}{\pi} \int_{-\infty}^{\infty} y(1 + t^2)/(y^2 + (x - t)^2) dE(t),$$
$$y > 0, k \geqq 0,$$

where $E(t)$ is a bounded increasing function on the closed infinite interval $[-\infty, \infty]$. We normalize $E(t)$ to be continuous from the right. This follows from a theorem of Herglotz [4] for $|w| < 1$ by the transformation $z = i(1 - w)/(1 + w)$. The term $-ky$ arises by subtracting a possible jump at infinity, so that $E(t) \to E(\infty)$ as $t \to \infty$, and $E(t) \to E(-\infty)$ as $t \to -\infty$. Let $u(x, y)$ denote the non-negative harmonic function given by the integral on the right hand side of (8). Then, by (7),

$$\int_{-\infty}^{\infty} u(x,y)/(1+x^2)\,dx \to 0 \quad \text{as } y \to 0+.$$

But

$$\int_{-\infty}^{\infty} u(x,y)/(1+x^2)dx = \tfrac{1}{\pi}\int_{-\infty}^{\infty}(1+x^2)^{-1}dx\int_{-\infty}^{\infty} y(1+t^2)(y^2+(x-t)^2)^{-1}dE(t)$$

$$= \tfrac{1}{\pi}\int_{-\infty}^{\infty}(1+t^2)dE(t)\int_{-\infty}^{\infty}\{(1+x^2)(y^2+(x-t)^2)\}^{-1}y\,dx$$

$$= \int_{-\infty}^{\infty}(1+y)(1+t^2)((1+y)^2+t^2)^{-1}\,dE(t) = \pi u(0,1+y),^{5}$$

where the interchange of the order of integration is valid on account of positivity. It follows that

$$0 = u(0,1) = \tfrac{1}{\pi}\int_{-\infty}^{\infty} dE(t) = \tfrac{1}{\pi}(E(\infty) - E(-\infty)).$$

Therefore, $E(t) \equiv$ constant, and $\log|\phi(x+iy)| = -ky$, $\phi(z) = e^{ikz+ic}$, where c is a real constant. This proves that $F(z) = e^{ikz+ic}b(z)$.

Remark 1. It is known [1] that for any $\theta_0 > 0$, $\log|b(z)|/|z| \to 0$ uniformly for $\theta_0 < \arg z < \pi - \theta_0$ as $|z| \to \infty$ outside a certain set $\Delta = \Delta(\theta_0)$ of finite logarithmic length.[6] Therefore, if in addition to the conditions stated in Theorem 2, one imposes the further requirement that for some $\theta (0 < \theta < \pi)$, $(\log|F(re^{i\theta})|)/r \to 0$ as $r \to \infty$, then $k = 0$ and $F(z) = e^{ic}b(z)$.

Remark 2. *A propos* the unit circle there is the following well-known and easily proved result:

Any holomorphic function $f(w)$ in $|w| < 1$ satisfying the conditions
(i) $|f(w)| < 1$,

(ii) $$\int_0^{2\pi} \log|f(re^{i\theta})|\,d\theta \to 0, \quad \text{as } r \to 1-,$$

is of the form $f(w) = e^{ic}b(w)$, where c is a real number and $b(w)$ is the Blaschke product formed with the zeros of $f(w)$.

Our uniqueness theorem is less precise than this in that the factor e^{ikz} can occur. Of course, this difference originates from the fact that (ii) does not correspond to (6).

MASSACHUSETTS INSTITUTE OF TECHNOLOGY.

---

[5] We use here the semi-group property of the kernel $c(x;\lambda,\mu) = \frac{1}{\pi}\lambda(\lambda^2+(x-\mu)^2)^{-1}$: $c(x;\lambda_1,\mu_1) * c(x;\lambda_2,\mu_2) = c(x;\lambda_1+\lambda_2,\mu_1+\mu_2)$, $*$ denoting convolution in $L(-\infty,\infty)$.

[6] The logarithmic length of a measurable subset Δ of $(0,\infty)$ is defined as $\int_\Delta dr/r$.

REFERENCES.

[1] R. P. Boas, Jr., *Entire Functions*, New York, 1954, p. 115.

[2] W. Kryloff, "On functions which are regular in a half-plane," *Recueil Mathematique*, vol. 6 (48) (1939), in particular pp. 113-116.

[3] A. Ostrowski, "Über die Bedeutung der Jensenschen Formel für einige Fragen der komplexen Funktionentheorie," *Acta Litterarum ac Scientiarum*, Szeged, 1923, pp. 80-81.

[4] M. H. Stone, *Linear Transformations in Hilbert Space*, American Mathematical Society Colloquium Publications, vol. XV (1932), pp. 570-571.

[5] E. C. Titchmarsh, *Theory of Functions*, 2nd ed., Oxford, 1939, p. 125.

SOLUTIONS OF SOME PROBLEMS OF DIVISION.*[1]

Part III. Division in the Spaces, $\mathfrak{D}'$, $\mathfrak{H}$, $\mathfrak{Q}_A$, $\mathfrak{O}$.

By Leon Ehrenpreis.

---

1. **Introduction.** The main division problem is the following: Let $\mathfrak{D}$ be a differential operator with constant coefficients and let T be a distribution (see [12]); can we find a distribution S which satisfies

$$DS = T \; ? \tag{1}$$

If T is a distribution of finite order, that is, if T can be written as a finite sum of derivatives of continuous functions, then the existence of an S satisfying (1) was proven by the author (see [2]) and, independently, by B. Malgrange in his thesis (see [11]). In the present paper we shall give the complete solution to the division problem. Moreover, we shall extend our results to the case where D is a partial differential-difference operator with constant coefficients. This general result is apparently new even in the case of ordinary differential-difference equations.

The question naturally arises: Which spaces of distributions or functions have the property that $f \rightarrow Df$ maps the space onto itself? Our main result is that the space $\mathfrak{D}'$ of all distributions has this property; previously we had shown that the space of distributions of finite order, and the space of indefinitely differentiable functions have this property.

For the space $\mathfrak{H}$ of entire functions (see [4]) even more is true: For any $W \in \mathfrak{H}'$, and any $f \in \mathfrak{H}$, we can find a $g \in \mathfrak{H}$ such that

$$W * g = f. \tag{2}$$

(This result was discovered independently by Malgrange in his thesis [11]). What other spaces have this property in addition to the space $\mathfrak{H}$?

We shall see that the solution of the two problems mentioned above can be translated into problems concerning the Fourier transform of the dual of the space in question. For example, the Fourier transform of the space $\mathfrak{D}$ is the space of all entire functions of exponential type which lie in the space

---

* Received August 28, 1955.

[1] Work supported by National Science Foundation Grant NSF5-G1010.

$\mathcal{S}$ of Schwartz (see [2], [12]), while the Fourier transform of the space $\mathcal{H}'$ is the space of all entire functions of exponential type. The question arises, what is the topology of these spaces of Fourier transforms? Once this topology is written down explicitly, it is not difficult to apply our previous methods (see [2], [3]) to show that $Dh \to h$ is a continuous linear map of $D\mathcal{D}$ into $\mathcal{D}$ ($D\mathcal{D}$ is the space of all Dh for $h \in \mathcal{D}$ with the topology induced by $\mathcal{D}$) and that $W * U \to U$ is a continuous linear map of $W * \mathcal{H}'$ into $\mathcal{H}'$. Once this is established, we can use the Hahn-Banach theorem to show that, for any $T \in \mathcal{D}'$ or for any $f \in \mathcal{H}$, equations (1) and (2) are solvable.

From the above, we see that there are two important steps in solving our main problems for a space A:

Problem 1. Characterize explicitly the Fourier transform of the dual A' of A and its topology.

Problem 2. Find a class of convolution maps $h \to W * h$ of $A \to A$ which have the property that $W * h \to h$ is a continuous map of $W * A'$ into A'.

Now, Problem 2 can usually be solved when Problem 1 can be solved and when the Fourier transform $\mathfrak{A}'$ of A' is a space of analytic functions. There are essentially two cases:

Case 1. The space $\mathfrak{A}'$ consists of functions defined only by regularity conditions or by growth conditions in the whole complex space. E.g. the space of all entire functions of exponential type, the space of all entire functions, the space of all entire functions of finite order. For these spaces we can use the minimum modulus results of [4] to show that, in most cases, if A' is a ring under convolution, then for every $W \in A'$, $W * h \to h$ is a continuous linear map of $W * A' \to A'$.

Case 2. The space $\mathfrak{A}'$ consists of functions defined by growth conditions in the whole complex space and additional conditions imposed on the real subspace. E.g. the Fourier transform of $\mathcal{D}$ or $\mathcal{E}'$ (see [5]), or many of the spaces which arise in the theory of infinite derivatives (see [7]). For these spaces, in order to prove that $W * h \to h$ is a continuous linear map of $W * A' \to A'$, we have to know that the Fourier transform of W does not decrease too rapidly on the real subspace. That is why, for these spaces, we can usually prove the continuity only for W a differential-difference

operator, where we can use the theory of mean-periodic functions (see [3], [4]) to get lower bounds on the real subspace.[2]

For many of the spaces we encounter, the solution of Problem 1 can be obtained by methods which are inspired by results in my paper on the theory of infinite derivatives (see [7]). This is not the case with the space $\mathfrak{D}$; the characterization of the topology of the Fourier transform $\boldsymbol{D}$ of $\mathfrak{D}$ is extremely difficult because the space $\mathfrak{D}$ cannot be defined by the methods of the theory of infinite derivatives. This is the reason why the solution of the division problem for $\mathfrak{D}'$ was so difficult to obtain.

Some of the results of this paper regarding the spaces $\mathfrak{D}$ and $\mathfrak{H}$ were announced in [8]. In addition, we shall show how to solve the division problems for the spaces

1. $\mathfrak{Q}_A$ of entire functions of order $\leqq A$;
2. $\mathfrak{O}$ of formal power series;
3. $\mathfrak{B}$ of convergent power series;
4. $\mathcal{E}$ of indefinitely differentiable functions.

Of course, our methods apply to many other spaces (see Section 7 below).

The notations of this paper will be that of Part I (see [2], Section 2) and the results of Parts I and II will be used here in an essential way (see [2], [3]).

2. Solution of the main division problem. As mentioned in the introduction, the first step in the solution of the division problem for $\mathfrak{D}'$ is the characterization of the topology of $\boldsymbol{D}$. Now, the characterization of the topology of $\boldsymbol{D}_F$ (see [2]) (for $n=1$) was obtained by means of giving bounds on functions of $\boldsymbol{D}_F$ on a sequence of lines parallel to the real axis. That we cannot define the topology of $\boldsymbol{D}$ in terms of bounds on horizontal lines is a consequence of a Phragmén-Lindelöf theorem (see [2], Lemma 1, p. 887). For, this shows that, if $F \in \boldsymbol{D}$, then a bound on Z^jF on any horizontal line implies a bound on the real axis. Thus, if a neighborhood N of zero in $\boldsymbol{D}$ is to be defined in terms of bounds of Z^jF on horizontal lines, only a finite number of j can occur. (For otherwise, we could find a $G \in \boldsymbol{D}$ so that, for no $a > 0$ is $aG \in N$.) Thus, we can only get neighborhoods of zero in $\boldsymbol{D}_F$ by means of bounds on horizontal lines.

[2] We can even ask the question of finding all W for which $W * h \to n$ is a continuous linear map of A' into A'. In case $A = \mathfrak{D}'$, this question has been solved by the author (see "Completely inversible operators," *Proceedings of the National Academy of Science*, vol. 41 (1955), pp. 945-946). The case $A = \mathcal{E}$ is also handled there.

Thus, if a "good" description of the topology of $\boldsymbol{D}$ is possible, it must be given in terms of bounds on something else. This "something else" which replaces horizontal lines is a sequence of curves:

THEOREM 1. *The topology of $\boldsymbol{D}$ can be described as follows: We choose first positive numbers c, η. Next, for each integer $l \geqq 0$, we choose*

(a) *Two non-negative integers a_l, d_l with $\{a_l\}$, $\{d_l\}$ strictly increasing to infinity, with $a_0 = d_0 = 0$, and $a_{l+1} \geqq a_l + c + 1$.*

(b) *A positive integer $b_l \geqq 5$ such that $\{b_l\}$ is monotonically increasing.*

We consider the set N of all $G \in \boldsymbol{D}$ which satisfy, for each l,

$$\max | \exp(3id_l p_k \cdot z) z_j^s G(z) | \leqq \eta, \text{ where } z \in \Gamma_l^k \tag{3}$$

for $s = 0, 1, 2, \cdots, b_l$, $j = 1, 2, \cdots, 2^n$, $k = 1, 2, \cdots, 2^n$, where Γ_0^k is, for each k, the set of z such that $|\mathfrak{I}(z_j)| \leqq c$ for all j, and Γ_l^k for $l > 0$ are defined as follows: For each $z \in C$, let γ_l be defined by $\gamma_0 = 0$ and, for $l > 0$,

$$\exp(d_l \gamma_l) = \begin{cases} |\zeta|^{b_l} & \text{for } |\zeta| \geqq 1 \\ 1 & \text{for } |\zeta| \leqq 1 \end{cases} \tag{4}$$

where $\zeta = (R(z_1), R(z_2), \cdots, R(z_n))$ and $|\zeta| = \sum |\zeta_j|$, that is,

$$\gamma_l = \begin{cases} b_l \log |\zeta| / d_l & \text{for } |\zeta| \geqq 1 \\ 0 & \text{for } |\zeta| \leqq 1. \end{cases} \tag{5}$$

Then Γ_l^k is the set consisting of all $z \in C^k$ for which

$$z_j = \zeta_j + i\gamma_l p_{k_j} + i p_{k_j} a_l + i\xi, \quad -c < \xi < c, \qquad j = 1, 2, \cdots, n. \tag{6}$$

ASSERTION. *The sets N described above form a fundamental system of neighborhoods of zero in $\boldsymbol{D}$.*

Proof. The sets N described above are obviously convex. Let $F \in \boldsymbol{D}$; then for some $\alpha > 0$, $F \in \boldsymbol{D}_\alpha$. Assume at first that l is chosen so large that $d_l \geqq \alpha$. Then, by a generalized Phragmén-Lindelöf theorem (see [2], Lemma 1, p. 887) for each k we know that $\exp(id_l p_k \cdot z) F(z)$ is bounded for $z \in C^k$ by the bound M of F on R. It follows that, whenever $d_l \geqq \alpha$, we have

$$\max_{z \in \Gamma_l^k} | \exp(3id_l p_k \cdot z) z_j^s F(z) | \leqq M \tag{7}$$

for $s = 0, 1, 2, \cdots, b_l$, $j = 1, 2, \cdots, n$, and $k = 1, 2, \cdots, 2^n$, because on Γ_l^k, $|\exp(id_l p_k \cdot z) z_j^s| \leqq 1$ for $s = 0, 1, 2, \cdots, b_l$ and for $j = 1, 2, \cdots, n$, as follows easily from the definitions.

Next we consider the finite number of l for which $d_l < \alpha$.

Let δ_l^k be the set of $z \in C^k$ of the form $z_j = \zeta_j + i\gamma_l p_{k_l} + i a_l p_{k_l}$ for all j; then for any $l \geqq 1$, on δ_l^k we have

$$(8) \qquad |\exp(-i d_l p_k \cdot z)| \leqq \exp(d_l a_l)(1 + \sum |z_j|^{n b_l}).$$

From (8) it follows that, on δ_l^k,

$$(9) \qquad |\exp(-i\alpha p_k \cdot z)| \leqq q_l(1 + \sum |z_j|^{\beta_l})$$

if β_l and q_l are chosen sufficiently large. Hence, for any $z \in \Gamma_l^k$, if z' denotes the point on δ_l^k such that $\mathcal{R}(z_j) = \mathcal{R}(z_j')$ for all j, we have by (9),

$$(10) \quad |\exp(-i\alpha p_k \cdot z)| \leqq \exp(\alpha) |\exp(-i\alpha p_k \cdot z')|$$
$$\leqq q_l(1 + \sum |z_j'|^{\beta_l}) \exp(c\alpha) \leqq q_l'(1 + \sum |z_j|^{\beta_l}) \text{ for some } q_l' > 0.$$

Now, let β_l' be an integer and let M' be a bound for $|z_j^{\beta_l'} F(z)|$, $|F(z)|$ on R for $j = 1, 2, \cdots, n$. By the Phragmén-Lindelöf theorem, M' is also a bound for $|\exp(i\alpha p_k \cdot z) z_j^{\beta_l'} F(z)|$ and $|\exp(i\alpha p_k \cdot z) F(z)|$ in C^k, *a fortiori* in Γ_l^k. Using (10) we deduce that, if β_l' is large enough, we can find an $a > 0$ so that aF satisfies (3) for this particular l. Since there are only a finite number of l for which $d_l < \alpha$, we can even choose a independent of l.

Summing up what we have done so far, we see:

(a) *For each $F \in \boldsymbol{D}$ we can find an $a > 0$ so that $aF \in N$.*

(b) *For each $\alpha > 0$ there exist positive numbers β'' and M'' so that the conditions $F \in \boldsymbol{D}_\alpha$,*

$$\max_{z \in R} |F(z)| \leqq M'', \qquad \max_{z \in R} |z_j^{\beta''} F(z)| \leqq M'' \text{ for } j = 1, 2, \cdots, n$$

imply $F \in N$.

By (a) it follows that the sets N define a locally convex topology λ on the set of functions in $\boldsymbol{D}$. By (b) together with Theorem 1 of [2] (see p. 887) it follows that for any $\alpha > 0$ the topology induced by λ on the set of functions of $\boldsymbol{D}_\alpha$ is weaker than (or the same as) the topology of the space $\boldsymbol{D}_\alpha$; trivial considerations show that this topology is the same as that of $\boldsymbol{D}_\alpha$. This proves (see [2]) that the topology λ is weaker than (or the same as) the topology of $\boldsymbol{D}$. In order to conclude the proof of Theorem 1, we must show that every neighborhood of zero in $\boldsymbol{D}$ contains some N. For this purpose, we shall need the following

LEMMA 1. *Let δ_l^k be the real n-chains defined as above. Then for any k, l, and any $F \in \boldsymbol{D}$, we have*

$$(11) \qquad \int_{\delta_l^k} F(z)\, dz = \int_R F(z)\, dz$$

where the integral on the left side of (11) *converges absolutely.* (*We also set* $\delta_0{}^k = R$ *for all* k.) *Moreover, for each* l *we can find a positive number* θ_l *which is independent of* a_l *so that for any continuous function* H *on* C, *the conditions*

$$\max_{z \in \delta_l{}^k} |z_j{}^s H(z)| \leqq \epsilon_l \tag{12}$$

for any k, *and for* $j = 1, 2, \cdots, n$, $s = 0, 1, 2, \cdots, 2n$, *imply*

$$\int_{\delta_l{}^k} |H(z)|\, dz \leqq 1. \tag{13}$$

Let us assume Lemma 1 for the present, and complete the proof of Theorem 1. Let Q be any neighborhood of zero in $\boldsymbol{D}$, and denote by $\tilde{Q}$ the set of inverse Fourier transforms of the functions of Q. By well-known properties of the topology of $\mathfrak{D}$, we may assume that $\tilde{Q}$ is given as follows: For each integer $r \geqq 0$, we are given a positive integer u_r and a positive number v_r, where we may assume $u_r \geqq u_{r-1}$ and $v_r \leqq v_{r-1}$ for all r. $\tilde{Q}$ consists of all $f \in \mathfrak{D}$ which satisfy

$$\max_{x \notin K_{12r(n+1)}} |((\partial^s/\partial x_j{}^s)f)(x)| \leqq v_r \tag{14}$$

for all r and for $j = 1, 2, \cdots, n$, $s = 0, 1, 2, \cdots, u_r$, and also

$$\max_{x \in R} |((\partial^s/\partial x_j{}^s)f)(x)| \leqq v_0 \tag{15}$$

for $j = 1, 2, \cdots, n$ and $s = 0, 1, 2, \cdots, u_0$.

We define the set N as follows: Let $c = \eta = 1$; set $d_l = l$; let $b_l = 2n + u_l$; next choose $a_l \geqq 1$ so that $e^{-a_l} \leqq v_l\theta_l$ and so that $\{a_l\}$ is monotonically increasing to infinity.

We claim that $N \subset Q$. First, it is clear that for any $G \in N$, the inverse Fourier transform g of G satisfies (15). Let $x \in R$, $x \notin K_{12l(n+1)}$. Let k be chosen so that $\operatorname{sgn} p_{k_j} = \operatorname{sgn} x_j$ for $j = 1, 2, \cdots, n$. By Lemma 1 we may write, for $s' = 0, 1, 2, \cdots, u_l$, $j = 1, 2, \cdots, n$,

$$\begin{aligned}(\partial^{s'}/\partial x_j{}^{s'} g)(x) &= \int_{\delta_l{}^k} i^{s'} z_j{}^{s'}\, G(z) \exp(ix \cdot z)\, dz \\ &= \int_{\delta_l{}^k} i^{s'} z_j{}^{s'} G(z) \exp(\tfrac{1}{2}(ix \cdot z)) \exp(\tfrac{1}{2}(ix \cdot z))\, dz.\end{aligned} \tag{16}$$

Since $x \notin K_{12l(n+1)}$ (we assume $l \geqq 1$), for at least one j, we have $|x_j| > 12(n+1)l$. Thus, for $z \in \delta_l{}^k$, $\mathfrak{J}(\tfrac{1}{2}x \cdot z) \geqq 6(n+1)la_l \geqq a_l$. This means that, for $z \in \delta_l{}^k$, we have

$$|\exp(\tfrac{1}{2}(ix \cdot z))| \leqq e^{-a_l} \leqq \theta_l v_l. \tag{17}$$

On the other hand, for $z \in \delta_l^k$, if $s'' = 0, 1, \cdots, u_l + 2n$, $j = 1, 2, \cdots, n$, we have

$$| z_j^{s''} G(z) \exp(\tfrac{1}{2}(ix \cdot z)) |$$
$$= | G(z) \exp(3ilp_k \cdot z) \exp(\tfrac{1}{2}(ix) \cdot z - 3ilp_k \cdot z) z_j^{s''} |.$$

For $z \in \delta_l^k$, $s'' = 0, 1, 2, \cdots, u_l + 2n$, $j = 1, 2, \cdots, n$, we have

$$| G(z) \exp(3ilp_k \cdot z) z_j^{s''} | \leqq 1,$$

by the definition of N. On the other hand, for $x \notin K_{12l(n+1)}$, for some j, $| x_j | \geqq 12l(n+1)$, so that, for $z \in \delta_l^k$,

$$| \mathfrak{I}(\tfrac{1}{2}x \cdot z - 3lp_k \cdot z) | \geqq 6l(n+1) | \mathfrak{I}(z_j) | - 3ln \, | \mathfrak{I}(z_j) | > 0$$

because, for any point z on δ_l^k, and any j', j'', $| \mathfrak{I}(z_{j'}) | = | \mathfrak{I}(z_{j''}) |$. Thus,

$$| z_j^{s''} G(z) \exp(\tfrac{1}{2}ix \cdot z) | \leqq 1. \tag{18}$$

Now, using (16), (17), (18) and Lemma 1, we deduce that, for $s' = 0, 1, 2, \cdots, u_l$, $j = 1, 2, \cdots, n$, $x \notin K_{12l(n+1)}$

$$| ((\partial^{s'}/\partial x_j^{s'}) g)(x) | \leqq v_l. \tag{19}$$

This means that $g \in \tilde{Q}$ so that $G \in Q$. This completes the proof that $N \subset Q$.

Proof of Lemma 1. We assume $k = 1$, and we write δ_l for δ_l^1; the case $k \neq 1$ is handled similarly. δ_l is (for z large) the set of $z \in C^1$ such that $\mathfrak{I}(z_j) = a_l + \gamma_l$, where $\gamma_l =$ constant $\log | \zeta |$, and where $\zeta = (R(z_1), \cdots, R(z_n))$. It follows easily that the element of surface area on δ_l is $\leqq$ constant $d\zeta$ where $d\zeta$ denotes the element of area on R (for $| \zeta |$ large). The existence of θ_l now results immediately.

In order to complete the proof of Lemma 1, we must show the following: For any $\psi > 0$, call V_ψ the real n-chain in C^1 consisting of all $z \in C^1$ such that $| \zeta | = \psi$, and $0 \leqq I(z_1) = I(z_2) = \cdots = I(z_n) \leqq \gamma_l$ for all j; we want to prove that, for any $F \in \boldsymbol{D}$, $\int_{V_\psi} F(z) \, dz \to 0$ as $\psi \to \infty$. Suppose that $F \in \boldsymbol{D}_\alpha$, and that $\gamma_l = a_l + \rho \log | \zeta |$ for $| \zeta |$ large, where ρ is some constant. Then we know that $\int_{V_\psi} dz = \rho' \psi^{n-1} (\log \psi + a_l)$ for ψ large, where ρ' is another constant.

We also know that we can find an $M > 0$ so that, for all $x \in R$, $j = 1, 2, \cdots, n$, we have $| x_j^{3n + [\alpha\rho] + 2} F(x) | \leqq M$ where $[\alpha\rho]$ is the greatest integer

$\leq x\rho$. Thus, by the Phragmén-Lindelöf theorem, for all $z \in V_\psi$, we have, for $j = 1, 2, \cdots, n$,

$$| x_j^{3n+[\alpha\rho]+2} F(z) | \leq M \exp(\alpha\gamma_!) \leq M' | \psi |^{\alpha\rho}$$

for some M'. It follows immediately that $\int_{V_\psi} | F(z) | dz \to 0$ which completes the proof of Lemma 1, and hence completes the proof of Theorem 1.

We have thus completed step (a) for the space $\mathfrak{D}'$ as explained in the introduction. We shall now pass to step (b). Since the functions of $\boldsymbol{D}$ are defined by both growth conditions in all of C and additional conditions on R, we should not expect that every distribution W which satisfies $W * \mathfrak{D} \subset \mathfrak{D}$ should also satisfy $W * \mathfrak{D}' = \mathfrak{D}'$. Indeed, if $W \in \mathfrak{D}$, then $W * \mathfrak{D}' \subset \mathcal{E}$, (see [12]) so that certainly, $W * \mathfrak{D}' \neq \mathfrak{D}'$. In this case it is not difficult to see that $W * \mathcal{E} \neq \mathcal{E}$ (see [11]). From this it follows that $W * W * \mathfrak{D}' \subset W * \mathcal{E} \neq \mathcal{E}$. (We have not been able to prove that $W * \mathfrak{D}' \neq \mathcal{E}$, although this seems to be certainly true.) On the other hand, if D is a punctual distribution (see [3]), i.e. D is a partial differential difference operator, then we can prove that $D * \mathfrak{D}' = \mathfrak{D}'$.

THEOREM 2. *Let P be an exponential polynomial, $P \neq 0$. Then $PG \to G$ is a continuous linear map of $P\boldsymbol{D} \to \boldsymbol{D}$.*

($P\boldsymbol{D}$ is the space of all PG, $G \in \boldsymbol{D}$ with the topology induced by $\boldsymbol{D}$. The properties of exponential polynomials which we shall use are derived in [3].)

Proof. We shall use induction on the number q of letters which occur in P. For $q = 0$, P is a non-zero constant and the result is obvious. Suppose $q > 0$ and that the theorem is proven for all exponential polynomials in fewer than q letters. Without loss in generality, we may assume that we may write

$$\begin{aligned} P(z) = [&P_0(z_2, \cdots, z_p) z_1^m \\ &+ P_1(z_2, \cdots, z_p) z_1^{m-1} + \cdots + P_m(z_2, \cdots, z_p)] \exp(ia_1 z_1) \\ &+ Q_2(z_1, \cdots, z_p) \exp(ia_2 z_1) + \cdots + Q_r(z_1, \cdots, z_p) \exp(ia_k z_1) \end{aligned}$$

where the P_j and the Q_j are exponential polynomials (the exponentials which occur in the Q_j being independent of z_1) with $P_0 \neq 0$, and where $a_1 < a_2 < \cdots < a_k$. The proof of Theorem 2 will be completed if we can show that, for some positive integer d, $PG \to P_0^d G$ is a continuous linear map of $P\boldsymbol{D}$ into $P_0^d \boldsymbol{D}$.

Let $L = a_k - a_1$. Let $z_2, \cdots, z_p$ be fixed and denote by P' the exponential polynomial in one variable: $z' \to P(z', z_2, \cdots, z_p)$. Let I be any interval of length $2(L+1)$ in the complex plane which is parallel to the real axis. By Lemma 1 of [3] we can find a point $z_0 \in I$ and constant K depending only on P such that $|P'(z_0)| \geqq K \exp(-|a_1 I(z_0)|) |P_0'|$, where $P_0' = P_0(z_2, \cdots, z_p)$. By Lemma 2 of [3] and the proof of Theorem 1 of [3] (see p. 288), we can describe about z_0 a circle of radius r' with $2(L+1) \leqq r' \leqq 4(L+1)$ such that, for all z' on this circle,

$$(20) \qquad |P'(z')| \geqq K'(1+|z'|)^B |P_0'|^{d'} \exp(-d' |a_1 I(z')|)$$

where B, d' and K' are constants which depend only on P, and d' is a positive integer. We set $d = d'$.

Let N be a neighborhood of zero in $\boldsymbol{D}$; let $m = 3 + d' |a_1|$. By Theorem 1 we may assume that N is defined by sequences $\{a_l\}$, $\{b_l\}$, $\{d_l\}$ and positive numbers c, η such that N consists of all $G \in \mathbf{D}$ which satisfy, for each l,

$$(21) \qquad \max |\exp(mid_l p_k \cdot z) z_j^s G(z)| \leqq \eta, \text{ where } z \in \Gamma_l^k,$$

for $k = 1, 2, \cdots, 2^n$, $j = 1, 2, \cdots, n$, and for $s = 0, 1, 2, \cdots, b_l$. (Theorem 1 is actually applicable only in case that m is replaced by 3, but it is clear from the proof of Theorem 1 that any number $\geqq 3$ will do.) We may also assume that $c \geqq 8(L+1)$.

For each l, k, let Θ_l^k be the region consisting of all $z \in C$ of the form

$$(22) \qquad z_j = z_j' + i\xi_j, \qquad z' \in \Gamma_l^k, \qquad |\xi_j| \leqq 8(L+1).$$

We consider the set N' of all $G \in \boldsymbol{D}$ which satisfy, for each l,

$$(23) \qquad \max |\exp(3id_l p_k \cdot z) z_j^s G(z)| \leqq \eta', \text{ where } z \in \Theta_l^k,$$

for $k = 1, 2, \cdots, 2^n$, $j = 1, 2, \cdots, n$, $s = 0, 1, 2, \cdots, b_l + [-B] + 1$, where $\eta' = \min(\eta/n, k'\eta/2n)(1 + 2^{-B})$ and where K', B are as in inequality (20). (If $B > 0$, then we can replace $[-B] + 1$ by 0.) Exactly as in the proof of Theorem 1 we can show that N' is a neighborhood of zero in $\boldsymbol{D}$.

We claim that the conditions $G \in \boldsymbol{D}$, $PG \in N'$ imply $P_0^d G \in N$. Let k, l be chosen; let z be any point in Γ_l^k. Let I be the interval in the complex plane, center z_1, length $2(L+1)$. We can find a point $z_0 \in I$ and a number r' with $2(L+1) \leqq r' \leqq 4(L+1)$ such that every point z' on the circle O, center z_0, radius r' satisfies (20). It is obvious from the definitions that, for any $z' \in O$, $(z', z_2, z_3, \cdots, z_n)$ lies in Θ_l^k. Thus, for any $z' \in O$, if $s = 0, 1, 2, \cdots, b_l$, $j = 1, 2, \cdots, n$, we have by (20),

$$(24)\quad |P_0'|^d\,|z_j'^s\,G(z', z_2, \cdots, z_n)\exp(mid_lp_k\cdot(z', z_2, z_3, \cdots, z_n))|$$
$$\leqq (1/K')\,|\exp(3id_lp_k\cdot(z', z_2, \cdots, z_n))\cdot z_j'^s|$$
$$\times\,|P(z', z_2, \cdots, z_n)(1+|z'|)^{-B}\,G(z', z_2, \cdots, z_n)|\leqq\eta$$

because $PG\in N'$, where we have written $z_j'=z_j$ for $j>1$, and $z_1'=z'$. Since the point z_1 is contained in O we also have, by the maximum modulus theorem,

$$(25)\qquad |P_0'|^d\,|z_j^s\,G(z_1, z_2, \cdots, z_n)\exp(mid_lp_k\cdot z)|\leqq\eta.$$

This proves that $P_0^dG\in N$.

We have shown that $PG\to P_0^dG$ is a continuous map of $P\boldsymbol{D}\to P_0^d\boldsymbol{D}$ at zero and hence, by linearity, everywhere. Since P_0^d is obviously an exponential polynomial in fewer than q letters, it follows from our induction assumption that $PG\to G$ is a continuous linear map of $P\boldsymbol{D}$ into $\boldsymbol{D}$ which is the desired result.

By means of Fourier transform we deduce immediately

COROLLARY. *For any partial differential-difference operator $D\neq 0$, $Dg\to g$ is a continuous linear map of $D\mathfrak{D}$ into $\mathfrak{D}$.*

THEOREM 3. *With D as above, for any $T\in\mathfrak{D}'$ we can find an $S\in\mathfrak{D}'$ such that $DS=T$.*

Proof. $Df\to T\cdot f$ is a continuous linear function $\tilde{S}$ on $D\boldsymbol{D}$ because it is the composition of the two continuous maps $Df\to f$ and $f\to T\cdot f$. By the Hahn-Banach theorem, $\tilde{S}$ can be extended to a continuous linear function on $\mathfrak{D}$, that is, $\tilde{S}$ can be extended to an $S\in\mathfrak{D}'$. For any $f\in\mathfrak{D}$ we have

$$DS\cdot f=S\cdot Df=\tilde{S}\cdot Df=T\cdot f.$$

This shows that $DS=T$ which is the desired result.

The question naturally arises as to what is the class of $W\in\mathcal{E}'$ which have the property that $W*\mathfrak{D}'=\mathfrak{D}'$. Theorem 3 shows that any punctual distribution has this property. By reasoning in a slightly more complicated manner than that used in the proof of Theorem 2, we can deduce

THEOREM 2′. *Let W be any distribution in $\mathcal{E}'$ of the form $W=D+h$ where D is a non-zero punctual distribution and h is a function of compact carrier which is sufficiently often differentiable.*[3] *Then $W*g\to g$ is a continuous linear map of $W*\mathfrak{D}\to\mathfrak{D}$.*

THEOREM 3′. *With W as above, given any $T\in\mathfrak{D}'$ we can find an $S\in\mathfrak{D}'$ which satisfies $W*S=T$.*

[3] We have not determined the best possible result.

3. Division in the space $\mathcal{E}$. We showed in a previous paper (see [13]) that every partial differential-difference operator maps $\mathcal{E}$ onto $\mathcal{E}$. We want to give a proof of this fact which is in keeping with the outline proposed in the introduction. For this purpose we must compute explicity the topology of the space $\boldsymbol{E}'$ (see [5]) which is the Fourier transform of the space $\mathcal{E}'$ of Schwartz; the characterization of this topology is itself of interest. (We write $|z| = \max(|z_j|.)$

THEOREM 4. *Let H be any continuous positive function on C such that, for any $k = 1, 2, \cdots, 2^n$ and for any $t > 0$*

$$(1 + |z|^t) \exp(-itp_p \cdot z) = O(H(z)). \tag{26}$$

Let N be the set of $F \in \boldsymbol{E}'$ such that, for any $z \in C$,

$$|F(z)| \leqq H(z). \tag{27}$$

Then the sets N form a fundamental system of neighborhoods of zero in $\boldsymbol{E}'$.

Proof. Let B be a bounded set in $\boldsymbol{E}'$. It is proven in [5] that we can find a positive number t' such that every $F \in B$ satisfies

$$|F(z)| \leqq t'(1 + |z|^{t'}) \exp(-it'p_k \cdot z)$$

for $z \in C^k$, $k = 1, 2, \cdots, 2^n$. Thus, given any set N as described in the statement of Theorem 4, we can find a positive number a such that $aB \subset N$, that is, N swallows every bounded set in $\boldsymbol{E}'$ (see [5]). Since the space $\boldsymbol{E}'$ is bornologic (see [7]) it follows that N is a neighborhood of zero in $\boldsymbol{E}'$.

Conversely, let Q be any neighborhood of zero in $\boldsymbol{E}'$. Now, (see [5]) the topology of $\boldsymbol{E}'$ can be described as follows: We consider each element of $\boldsymbol{E}'$ as defining, by multiplication, a continuous linear map of $\boldsymbol{D}_F$ into $\boldsymbol{D}_F$; then the topology of $\boldsymbol{E}'$ is that obtained by giving this set of maps the compact-open topology. We want to produce a neighborhood of zero N in $\boldsymbol{E}'$ of the type described in the statement of Theorem 4 such that $N \subset Q$. By the above, it is sufficient to show that, if B is any bounded set in $\boldsymbol{D}$, and if M is a neighborhood of zero in $\boldsymbol{D}_F$, we can find a neighborhood of zero N in $\boldsymbol{E}'$, as above, such that the conditions $F \in B$, $G \in N$ imply $FG \in M$.

Now, M may be described as follows (see [7]; the result may also be deduced without much difficulty from Theorem 7 of [2], p. 893): Let H_1 be a continuous function on R with the property: For each $t > 0$, $\exp(t|x|) = O(H_1(x))$; let m be a given integer. Then M contains in the set of all $F \in \boldsymbol{D}_F$ for which

(28) $\max |z_j{}^s F(z)| \leqq H_1(y)$, where
$$\mathfrak{J}(z_1) = y_1, \mathfrak{J}(z_2) = y_2, \cdots, \mathfrak{J}(z_n) = y_n,$$

with $y = (y_1, y_2, \cdots, y_n)$, for $s = 0, 1, \cdots, m$ and for $j = 1, 2, \cdots, n$.

B may be described as follows (the proof is essentially the same as the proof of Theorem 1 of [2], p. 887): For each integer s we are given a positive number $A_s \geqq s!$; we are also given a positive number A. B is contained in the set of all $F \in \boldsymbol{D}_F$ which are of exponential type $\leqq A$ and satisfy, for all s,

$$\max_{x \in R} |z_j{}^s F(x)| \leqq A_s; \qquad j = 1, 2, \cdots, n. \tag{29}$$

Now, any $F \in B$ satisfies (29) and so also satisfies, by the Phragmén-Lindelöf theorem,

$$\max_{z \in C^k} |z_j{}^s F(z) \exp(iAp_k \cdot z)| \leqq A_s \tag{30}$$

for $j = 1, 2, \cdots, n$, $k = 1, 2, \cdots, 2^n$, and all s.

Define the function H_2 on R by

$$H_2(x) = \sum_{r=0}^{\infty} \sum_{j=1}^{n} |x_j|^r / 2^{r+1} n A_r. \tag{31}$$

Since $A_r \geqq r!$, this series converges for all r. Then (31) and (29) imply that, for any $x \in R$, $F \in B$, we have $|F(x) H_2(x)| \leqq 1$. We set

$$H(z) = H_2(x) H_1(y) |\exp(iAp_k \cdot z)| / (1 + |z|^m) \tag{32}$$

for any $z \in C^k$, $k = 1, 2, \cdots, 2^n$, where

$$x = (\mathfrak{R}(z_1), \cdots, \mathfrak{R}(z_n)), \qquad y = (\mathfrak{J}(z_1), \cdots, \mathfrak{J}(z_n)).$$

It is clear from (30), (31), and (32) that, for any $z \in C$, for any $j = 1, 2, \cdots n$, and for any $F \in B$, we have

$$|z_j{}^s H(z) F(z)| \leqq H_1(y) \tag{33}$$

for $s = 0, 1, \cdots, m$. Using (33) we deduce that the set N of $G \in \boldsymbol{D}_F$ which satisfy $|G(z)| \leqq H(z)$ for all $z \in C$ has the property that $NB \in M$. Since H clearly satisfies the relation (23), this completes the proof of Theorem 4.

THEOREM 5. *Let $W \in \mathcal{E}'$ be of the form $W = D + f$ where D is a non-zero partial differential-difference operator with constant coefficients (punctual distribution) and $f \in \mathcal{E}'$ is a sufficiently often differentiable function.*[4] *Then $W * U \to U$ is a continuous linear map of $W * \mathcal{E}'$ into $\mathcal{E}'$.*

---

[4] The same remarks apply as in footnote 2.

The proof of Theorem 5 proceeds along essentially the same lines as the proof of Theorem 2 (and is, in fact, somewhat easier); we shall omit the details.

By the methods of proof of Theorem 3 we can now deduce (see [3])

THEOREM 6. *With W as above, for any $f \in E$ there is a $g \in E$ such that $W * g = f$.*

4. Division in the space H. By $\mathfrak{H}$ we denote the space of entire functions on C with its usual topology (see [4]). Each element of $\mathfrak{H}'$ can be represented as a distribution of compact carrier on C. The Fourier transform $\boldsymbol{H}'$ of $\mathfrak{H}'$ is the space of all entire functions of exponential type on C with no extra conditions on R. As mentioned in the introduction, we should therefore expect that every $W \in \mathfrak{H}'$ has the property that $W * U \to U$ is a continuous linear map of $W * \mathfrak{H}'$ into $\mathfrak{H}'$. That this is actually the case will be seen later (see the corollary to Theorem 8 below). First we shall have to describe the topology of $\boldsymbol{H}'$ explictly:

THEOREM 7. *Let H be any continuous positive function on C such that, for any $t > 0$,*

$$\exp(t\,|\,z\,|) = O(H(z)). \tag{34}$$

We consider the set N of $F \in H'$ such that $|F(z)| \leqq H(z)$. Then these sets N form a fundamental system of neighborhoods of zero in $\boldsymbol{H}'$.

Proof. First, let N be a set satisfying the hypotheses of Theorem 7. Then N is clearly convex. If B is any bounded set in $\boldsymbol{H}'$ then (the proof is similar to that of Proposition 2 of [4], p. 296) we can find positive numbers b, c so that all $F \in B$ satisfy $|F(z)| \leqq b \exp(c\,|\,z\,|)$. Thus, by (34), we can find an $a > 0$ such that $aB \subset N$, that is, N swallows every bounded set in $\boldsymbol{H}'$ (see [5]). Now, $\boldsymbol{H}'$ is bornologic (see [9]); this means that N is a neighborhood of zero in $\boldsymbol{H}'$.

(We can also prove that N is a neighborhood of zero in H' by using the method of proof of Theorem 14 below, without using the results of Grothendieck. This method is of use when we do not know in advance that the space in question is bornologic. Actually, we could deduce in this manner that $\boldsymbol{H}'$ is bornologic.)

Conversely, let M be a neighborhood of zero in $\boldsymbol{H}'$; we want to produce a neighborhood of zero N of the type described in the statement of Theorem 7 such that $N \subset M$. Call Q the set of inverse Fourier transforms of the functions of M (see [4]). By the definition of the topology of a dual space, we

can find a bounded set K in $\mathfrak{H}$ so that the conditions $U \in \mathfrak{H}'$, $|U \cdot f| \leqq 1$ for all $f \in K$ imply $U \in Q$.

Now (see [4], Proposition 1 and Equation (1) on p. 296), if we write $G \in \boldsymbol{H}'$ in the form

$$G(z) = \sum (i^{k_1+k_2+\cdots+k_n}/k_1!k_2!\cdots k!_n)\ G_{k_1k_2\cdots k_n} z_1^{k_1} z_2^{k_2}\cdots z_n^{k_n}$$

and if we write $f \in \mathfrak{H}$ in the form $f(z) = \sum f_{k_1k_2\cdots k_n} z_1^{k_1} z_2^{k_2}\cdots z_n^{k_n}$ then we have

$$F^{-1}G \cdot f = \sum G_{k_1k_2\cdots k_n} f_{k_1k_2\cdots k_n}. \tag{35}$$

Since K is bounded in $\mathfrak{H}$, for each system of positive integers l we can find a positive number $c_{l_1l_2\cdots l_n}$ such that every $f \in K$ satisfies

$$\max |f(z)| \leqq c_{l_1l_2\cdots l_n}, \text{ where } |z_1| \leqq l_1, |z_2| \leqq l_2, \cdots, |z_n| \leqq l_n. \tag{36}$$

From (36) and Cauchy's formula we deduce that the inequality

$$|f_{k_1k_2\cdots k_n}| \leqq c_{l_1l_2\cdots l_n}/l_1^{k_1} l_2^{k_2}\cdots l_n^{k_n} \tag{37}$$

holds for all $f \in K$ and for all integers $k_1, k_2, \cdots, k_n$, $l_1, l_2, \cdots, l_n$. We may clearly assume that $c_{11\cdots 1} = 1$ and that $c_{l_1l_2\cdots l_n} \leqq c_{l_1'l_2'\cdots l_n'}$ if

$$l_1 \leqq l_1', l_2 \leqq l_2', \cdots, l_n \leqq l_n'.$$

Let $\{\gamma_j\}$ be a strictly increasing sequence of numbers with $\gamma_0 = 0$. Let $\{d_j\}$ be a sequence of non-negative numbers with $d_0 = 0$, $d_j > 0$ for $j > 0$. We require further that

$$d_j + \gamma_j \leqq \gamma_{j+1} \tag{a}$$

for any j. We call $\alpha_{j_1j_2\cdots j_n}$ the region consisting of all $y \in C$ for which $d_{j_k} + \gamma_{j_k} \leqq |y_k| \leqq \gamma_{j_k+1}$.

$$\begin{aligned} \exp[\tfrac{1}{2}(j_1|y_1| + j_2|y_2| + \cdots + j_n|y_n|)]/c_{j_1+1,\cdots,j_n+1} \\ \geqq \exp[\tfrac{1}{2}((j_1-1)|y_1| + (j_2-1)|y_2| + \cdots + (j_n-1)|y_n|)], \end{aligned} \tag{b}$$

whenever $j_1 \geqq 1, j_2 \geqq 1, \cdots, j_n \geqq 1$, and whenever $|y_k| \geqq d_{j_k} + \gamma_{j_k}$ for all k.

(c) There is no point of the form l/m in the interval $\gamma_j \leqq x \leqq \gamma_j + d_j$ for any j, where m, l are integers and $m \leqq 2(j+1)$.

The existence of the sequences $\{\gamma_j\}$, $\{d_j\}$ is obvious. Moreover, it is clear that the regions $\alpha_{j_1,j_2,\cdots,j_n}$ do not overlap.

We set

$$c_{j_1+1\ j_2+1\ \cdots\ j_n+1}\ H(z) = \exp(\tfrac{1}{2}(j_1|z_1| + j_2|z_2| + \cdots + j_n|z_n|)), \quad z \in \alpha_{j_1j_2\cdots j_n} \tag{38}$$

and H, is defined so as to be continuous and monotonic (i.e. $H(z) \leqq H(z')$ if $|z_j| \leqq |z_j'|$ for $j=1,2,\cdots,n$). It is obvious from (38) and (b) above that H satisfies (34).

Denote by N the set of all $G \in \boldsymbol{H}'$ for which $|G(z)| \leqq H(z)$ for all $z \in C$. Let $G \in N$ and let $(k_1, k_2, \cdots, k_n)$ be an n-tuple of non-negative integers. It is not difficult to see by (c) that we can find an n-tuple of integers $(l_1, l_2, \cdots, l_n)$ and an n-tuple of positive numbers $(a_1, a_2, \cdots, a_n)$ such that

$$(39) \quad k_1/l_1 \geqq a_1 \geqq k_1/l_1 + 1, \qquad k_2/l_2 \geqq a_2 \geqq k_2/l_2 + 1, \cdots, k_n/l_n \geqq a_n \geqq k_n/l_n + 1$$

and such that the n-chain Ξ in C defined by $|z_1| = a_1$, $|z_2| = a_2, \cdots, |z_n| = a_n$ lies in $\alpha_{l_1 l_2 \cdots l_n}$.

By Cauchy's formula we may write[5]

$$G_{k_1 k_2 \cdots k_n} = -ik_1!k_2! \cdots k_n! \int_{\Xi} G(z)/z_1^{k_1+1} z_2^{k_2+1} \cdots z_n^{k_n+1}\, dz.$$

Now, on Ξ, we have

$$|G(z)| \leqq H(z) = \exp(\tfrac{1}{2}(a_1 l_1 + a_2 l_2 + \cdots + a_n l_n))/c_{l_1+1, l_2+1, \cdots, l_n+1}$$

by (38). Thus, on making use of (39) we find

$$c_{l_1+1, l_2+1, \cdots, l_n+1} k_1^{k_1} k_2^{k_2} \cdots k_n^{k_n} |G_{k_1 k_2 \cdots k_n}|$$
$$\leqq k_1!k_2! \cdots k_n! \exp(\tfrac{1}{2}(k_1 + k_2 + \cdots + k_n))(l_1+1)^{k_1}(l_2+1)^{k_2} \cdots (l_n+1)^{k_n}.$$

By Stirling's formula, we can find a constant θ such that

$$(40) \quad c_{l_1+1, l_2+1, \cdots, l_n+1} |G_{k_1 k_2 \cdots k_n}|$$
$$\leqq \theta \exp(-\tfrac{1}{2}(k_1 + k_2 + \cdots + k_n))(l_1+1)^{k_1}(l_2+1)^{k_2} \cdots (l_n+1)^{k_n}.$$

Using (35), (37), and (40) it is obvious that we can find a $\theta' > 0$ such that every $G \in N$ satisfies

$$(41) \qquad F^{-1} G \cdot f \leqq \theta' \text{ for all } f \in K.$$

This proves that $(1/\theta')N \subset M$ which completes the proof of Theorem 7.

THEOREM 8. *Let $J \in \boldsymbol{H}'$, $J \neq 0$; then $JG \to G$ is a continuous linear map of $J\boldsymbol{H}'$ into $\boldsymbol{H}'$.*

---

[5] We normalize the measure in such a manner that the usual factor $(2\pi)^n$ does not appear.

Proof. The idea of the proof is to use the following minimum modulus theorem for entire functions of exponential type (see [4], Theorem 5, p. 317): Suppose k is an entire function of exponential type of one complex variable and that $|k(z)| \leqq M \exp(A|z|)$ for all z. Then for any $r > 0$ there is an r' with $r \leqq r' \leqq 2r$ such that

$$\min_{|z|=r'} |k(z)| \geqq |k(0)|^d \exp(B \log M + cAr') \tag{42}$$

where B and c are certain constants and d is a positive integer.

We shall prove Theorem 8 by induction on the number q of variables on which J depends. For $q = 0$, J is a non-zero constant and the result is obvious; we assume $q > 0$ and that the result is known for entire functions of exponential type in fewer than q variables. We may clearly assume that J depends on $Z_1, Z_2, \cdots, Z_q$. For any complex numbers $z_2, \cdots, z_q$, set $J_1(z_2, \cdots, z_q) = J(0, z_1, \cdots, z_q)$. It is sufficient to show that $JG \to J_1{}^d G$ is a continuous linear map of $J\boldsymbol{H}' \to J_1{}^d\boldsymbol{H}'$.

By linearity, we must show the following: Given any neighborhood of zero N in $\boldsymbol{H}'$ there exists a neighborhood of zero N_1 in $\boldsymbol{H}'$ such that the conditions $G \in \boldsymbol{H}'$, $JG \in N_1$ imply $J_1{}^d G \in N$. By Theorem 6 we may assume that N consists of all $F \in \boldsymbol{H}'$ for which $|F(z)| \leqq H(z)$ for all $z \in C$, where H is a function satisfying (34). We may clearly assume that H is monotonic, i. e.

$$H(z') \geqq H(z) \text{ if } |z_1'| \geqq |z_1|, |z_2'| \geqq |z_2|, \cdots, |z_n'| \geqq |z_n|. \tag{43}$$

Now, we can clearly find positive numbers A, M such that $|J(z)| \leqq M \exp(A|z|)$ for all $z \in C$. We define the function H_1 on C by

$$H_1(2z) = M^B \exp(2cA|z_1| + BA|z_2| + \cdots + BA|z_q|) H(z_1, z_2, \cdots, z_n) \tag{44}$$

where B and c are the constants that appear in (42). It is clear that H_1 satisfies (34). Denote by N_1 the neighborhood of zero in $\boldsymbol{H}'$ consisting of all $G \in \boldsymbol{H}'$ for which $|G(z)| \leqq H_1(z)$ for all $z \in C$. We claim that the conditions $JG \in N_1$ implies $J_1{}^d G \in N$.

For, given any $z \in C$, we can find by (42) a number r' with $|z_1| \leqq r' \leqq 2|z_1|$ and

$$\min_{|z_1'|=r'} J(z_1', z_2, \cdots, z_q) \geqq |J_1(z_2, \cdots, z_q)|^d M^B \exp(cAr' + BA|z_2| + \cdots + BA|z_q|).$$

Therefore, if $|z_1'| = r'$, and $JG \in N_1$, we have

$$\begin{aligned} |J_1(z_2,\cdots,z_q)|^d\,|G(z_1',z_2,\cdots,z_n)| \\ \leqq M^{-B}\exp(-cAr' - BA|z_2| - \cdots - BA|z_q|) \\ \times |J(z_1',z_2,\cdots,z_q)G((z_1',z_2,\cdots,z_n)| \\ \leqq M^{-B}\exp(-cAr' - BA|z_2| - \cdots - BA|z_q|)H_1(z_1',z_2,\cdots,z_n). \end{aligned}$$

Since $|z_1| \leqq r' \leqq 2|z_1|$ we have, by the maximum modulus theorem,

$$\begin{aligned} |J(z_2,\cdots,z_q)|^d\,|G(z_1,z_2,\cdots,z_n)| \\ \leqq M^{-B}\exp(-2cA|z_1| - BA|z_2| - \cdots - BA|z_q|) \\ \times H_1(2z_1,2z_2,\cdots,2z_n) = H(z_1,z_2,\cdots,z_n) \end{aligned}$$

because (see (43) and (44)) H_1 is monotonic, and $c < 0$. This completes the proof of Theorem 8.

By taking the Fourier transform we deduce

COROLLARY. *For any $W \in \mathfrak{H}'$, $W \neq 0$, $W * U \to U$ is a continuous linear map of $W * \mathfrak{H}'$ into $\mathfrak{H}'$.*

THEOREM 9. *For any $W \in \mathfrak{H}'$, $f \to W * f$ maps $\mathfrak{H}$ onto $\mathfrak{H}$ continuously.*

Theorem 9 follows from the corollary to Theorem 8 exactly in the same manner as Theorem 3 is deduced from the Corollary to Theorem 2.

5. Entire functions of finite order. Let f be an entire function, and let $A \geqq 1$. We say that f is of order $\leqq A$ if, for every $\epsilon > 0$, we have

$$f(z) = O(\exp|z|^{A+\epsilon}). \tag{45}$$

The greatest lower bound of A for which (45) holds is called the *order* of f. By $\mathfrak{Q}_A$ we denote the space of all entire functions of order $\leqq A$; it is clear that $\mathfrak{Q}_A$ is a ring under multiplication. We define a topology in $\mathfrak{Q}_A$ by means of the semi-norms

$$v_\epsilon(f) = \sup_{z \in C}\exp(-|z|^{A+\epsilon})|f(z)| \tag{46}$$

for any $f \in \mathfrak{Q}_A$.

PROPOSITION 1. *$\mathfrak{Q}_A$ is a complete metrizable locally convex topological vector space,* i.e. *$\mathfrak{Q}_A$ is a Frechet space. $\mathfrak{Q}_A$ is not a Banach space. $\mathfrak{Q}_A$ is a Montel space, so that $\mathfrak{Q}_A$ is reflexive. $\mathfrak{Q}_A$ is also a Schwartz space.*[6]

[6] The theory of Schwartz spaces has been developed by Grothendieck in [9].

Proof. It is obvious that $\mathfrak{Q}_A$ is a locally convex topological vector space. Moreover, we may clearly define the topology of $\mathfrak{Q}_A$ by means of the semi-norms $v_{1/m}$ for m integral; thus, $\mathfrak{Q}_A$ is metrizable. In order to show completeness, we may therefore restrict ourselves to sequences. If $\{f_j\}$ is a Cauchy sequence in $\mathfrak{Q}_A$, we know, from the fact that $\mathfrak{H}$ is a Frechet space (see [4]), that $\{f_j\}$ converges to a function $g \in \mathfrak{H}$ on the topology of $\mathfrak{H}$. It is clear that, for every $\epsilon > 0$, $g(z) = O(\exp(|z|^{A+\epsilon})$ so that $g \in \mathfrak{Q}_A$. It is readily verified that $\{f_j\}$ converges to g in the topology of $\mathfrak{Q}_A$. It is clear that no neighborhood of zero in $\mathfrak{Q}_A$ is bounded, so $\mathfrak{Q}_A$ is not a Banach space.

We shall show now that $\mathfrak{Q}_A$ is a Montel space, that is, the closed bounded sets of $\mathfrak{Q}_A$ are compact. We shall show even more: For each $\epsilon > 0$ the set N_ϵ of $f \in \mathfrak{Q}_A$ which satisfy $v_\epsilon(f) \leqq 1$ is precompact for the topology induced by $v_{\epsilon'}$ whenever $\epsilon' > \epsilon$; this implies that $\mathfrak{Q}_A$ is a Schwartz space. Let $\{f_j\}$ be a sequence in N_ϵ; we can extract a subsequence $\{f_j\}$ which converges in $\mathfrak{H}$, say $f_{j_k} \to f$ in $\mathfrak{H}$, because $\mathfrak{H}$ is a Montel space and N_ϵ is clearly bounded in $\mathfrak{H}$; it is obvious that $f \in N_\epsilon$. Since each $f_{j_k} \in N_\epsilon$ we deduce immediately that $v_{\epsilon'}(f_{j_k} - f) \to 0$. This completes the proof of Proposition 1.

PROPOSITION 2. *For any $f \in \mathfrak{H}$, write* $f(z) = \sum f_{k_1 k_2 \cdots k_n} z_1^{k_1} z_2^{k_2} \cdots z_n^{k_n}$. *Then $f \in \mathfrak{Q}_A$ if and only if*

$$(47) \quad \liminf \log(1/|f_{k_1 k_2 \cdots k_n}|)/\{(k_1 \log k_1)(k_2 \log k_2) \cdots (k_n \log k_n)\} = 1/A.$$

In case $n = 1$ this Proposition is well-known (see [13], p. 253); the proof for $n > 1$ can be accomplished by a similar method.

By $\mathfrak{Q}_A'$ we denote the dual of $\mathfrak{Q}_A$ with the topology of uniform convergence on the compact (bounded) sets of $\mathfrak{Q}_A$. For any $U \in \mathfrak{Q}_A'$ and any integers $k_1, k_2, \cdots, k_n$ we set

$$(48) \qquad U_{k_1 k_2 \cdots k_n} = U \cdot Z_1^{k_1} Z_2^{k_2} \cdots Z_n^{k_n}.[7]$$

It is not difficult to show

PROPOSITION 3. *For any $f \in \mathfrak{Q}_A$, $\sum f_{k_1 k_2 \cdots k_n} Z_1^{k_1} Z_2^{k_2} \cdots Z_n^{k_n}$ converges to f in the topology of $\mathfrak{Q}_A$. Thus, for any $U \in \mathfrak{Q}_A'$,*

$$(49) \qquad U \cdot f = \sum U_{k_1 k_2 \cdots k_n} f_{k_1 k_2 \cdots k_n}$$

where the series on the right side of (49) *converges absolutely.*

For any $U \in \mathfrak{Q}_A'$, we can find an $\epsilon > 0$ so that U is bounded on the set of $f \in \mathfrak{Q}_A$ for which $v_\epsilon(f) \leqq 1$. By the Hahn-Banach theorem, U can

[7] For each j, Z_j is the function on C: $z \to z_j$.

be extended to a continuous linear function on the space $V_{A+\epsilon}$ of all continuous functions on C which are $o(\exp(|z|^{A+\epsilon})$. Thus, we have

PROPOSITION 4. *For each $U \in \mathfrak{Q}_A'$ we can find an $\epsilon > 0$ so that U can be represented by a measure which we again denote by U which falls off like* $\exp(-|z|^{A+\epsilon})$, that is,

$$(50) \qquad \int_{|z| \leqq r} d|U(z)| = O(\exp(-r^{A+\epsilon})).$$

If B is a bounded set in $\mathfrak{Q}_A'$, then the ϵ and symbol "O" in (50) *can be chosen uniformly for $U \in B$.*

For any $U \in \mathfrak{Q}_A$, we define the *Fourier transform* of U as the function on C:

$$(51) \qquad f \to U \cdot \exp(iz \cdot) = \int \exp(iz \cdot t) dU(t) = (\boldsymbol{F}(U))(z).$$

It follows immediately from Proposition 4 that $\boldsymbol{F}(U)$ is an entire function.

We define A' by $1/A + 1/A' = 1$, that is, A' is the conjugate exponent of A (see [15]). We wish to recall Hölder's inequality (see [13], p. 383 inequality (5)): For any complex numbers x, y we have

$$(52) \qquad |x|^A/A + |y|^{A'}/A' \geqq |xy|.$$

Using the fact that for any z, $t \in C$ we have $|z \cdot t| \leqq |z|\,|t|$, we deduce from (52)

$$(52') \qquad |z|^A/A + |t|^{A'}/A' \geqq |z \cdot t| \text{ for any } z, t \in C.$$

PROPOSITION 5. *For any $U \in \mathfrak{Q}_A'$, we can find an $\epsilon' > 0$ such that $\boldsymbol{F}(U)$ is an entire function of order $\leqq A' - \epsilon'$.*[8] *If B is a bounded set in $\mathfrak{Q}_A'$ we can choose ϵ' uniformly for $U \in B$. We may write:*

$$(53) \qquad (\boldsymbol{F}(U))(z) = \sum (i^{k_1+k_2+\cdots+k_n}/k_1!k_2!\cdots k_n!)(U_{k_1k_2\cdots k_n} z_1^{k_1} z_2^{k_2}\cdots z_n^{k_n}).$$

Proof. We apply Proposition 3 to equation (51), using the fact that

$$\exp(iz \cdot) = \sum (i^{k_1+k_2+\cdots+k_n}/k_1!k_2!\cdots k_n!)(z_1Z_1)^{k_1}(z_2Z_2)^{k_2}\cdots(z_nZ_n)^{k_n}$$

and we deduce (53) immediately. The first part of Proposition 5 can now be deduced from the fact that the numbers $U_{k_1k_2\cdots k_n}$ render the right side of (49) convergent whenever $f \in \mathfrak{Q}_A$ if and only if, for some $\epsilon' > 0$,

$$\sum (i^{k_1+k_2+\cdots+k_n}/k_1!k_2!\cdots k_n!) U_{k_1k_2\cdots k_n} Z_1^{k_1} Z_2^{k_2}\cdots Z_n^{k_n}$$

[8] In case $A = 1$, $A' = \infty$, and $A' - \epsilon'$ is to be interpreted as some (finite) number. A similar convention is to be employed throughout the rest of this paper.

defines an entire function of order $\leqq A'-\epsilon'$. (This fact can be proved by the methods of [7].) The part of Proposition 5 referring to the bounded set B can be handled similarly.

Another proof of the first part of Proposition 5 can be obtained from Hölder's inequality (52′), and Proposition 4. Thus, if ϵ' is determined by the condition $1/(A+\epsilon)+1/(A'-\epsilon')=1$, then we have, for any $z \in C$,

$$(54) \qquad (\boldsymbol{F}(U))(z)=\int \exp(iz\cdot t)dU(t) \leqq \int \exp(|z\cdot t|d|U(t)|$$
$$\leqq \int \exp(|z|^{A'-\epsilon'}/(A'-\epsilon')+t^{A-\epsilon}/(A+\epsilon))d|U(t)|$$
$$=\exp(z^{A'-\epsilon'}/(A'-\epsilon'))\int \exp(t^{A+\epsilon}/(A+\epsilon))d|U(t)|=O(\exp(z^{A'-\epsilon'})).$$

Similarly, if U lies in a bounded set b in $\mathfrak{Q}_A'$, the symbol "O" in (54) can be chosen uniformly for $U \in B$. This completes the proof of Proposition 5.

By means of Hölder's inequality (52′) we can also establish easily

PROPOSITION 6. *For each $\epsilon>0$ there exists a constant $c_\epsilon>0$ such that, for all $z \in C$,*

$$(55) \qquad v_\epsilon(\exp(iz\cdot)) \leqq c_\epsilon |z|^{A'-\epsilon'}$$

where, $1/(A+\epsilon)+(A'-\epsilon')=1$.

Remark. Proposition 6 furnishes us with another proof of the first part of Proposition 5.

We now want to derive the converse of Proposition 5, that is, $\boldsymbol{Q}_A'=\boldsymbol{F}(\mathfrak{Q}_A')$ consists of all entire functions of order $<A'$. We also want to describe completely the topology of $\boldsymbol{Q}_A'$ in order to study the division problem in the space $\mathfrak{Q}_A$. ($\boldsymbol{Q}_A'$ is given the topology to make $\boldsymbol{F}$ a topological isomorphism.)

PROPOSITION 7. *Let B be a set in $\boldsymbol{Q}_A'$; then B is bounded in $\boldsymbol{Q}_A'$ if and only if, for some $\epsilon'>0$ we can find an $M>0$ such that*

$$|G(z)| \leqq M\exp(|z|^{A'-\epsilon'}) \ \textit{for all}\ z\in C,\ G \in B.$$

Proof. The necessity of the condition was established in Proposition 5. Suppose then that B satisfies the stated condition; let ϵ be defined by $1/(A+\epsilon)+1/(A'-\epsilon')=1$. (We assume, as we may, that $A'-\epsilon'>1$.) We claim that, for some $M'>0$, $|\boldsymbol{F}^{-1}G\cdot f|\leqq 1$ for all $f \in N_{\epsilon/2}$. This is, in fact, an easy consequence of Proposition 3 and Cauchy's formula.

PROPOSITION 8. *Let* $\{G^j\}$ *be a sequence in* $\boldsymbol{Q}_A'$. *A necessary and sufficient condition that* $G^j \to 0$ *in* $\boldsymbol{Q}_A'$ *is that there exist an* $\epsilon' > 0$ *such that*

$$(56) \qquad \max_{z \in C} \exp(-|z|^{A'-\epsilon'} G^j(z)) \to 0.$$

For each j, *denote by* U^j *the inverse Fourier transform of* G^j. *Then if* (56) *is satisfied, we even have* $U^j \cdot f \to 0$ *uniformly on the set of* $f \in N_{\epsilon/2}$.

Proof. Suppose first that $G^j \to 0$ in $\boldsymbol{Q}_A'$. Then it follows from Proposition 1 that, for some $\epsilon_1 > 0$, $U^j \cdot f \to 0$ uniformly for $f \in N_{\epsilon_1}$. (See [9]; see also [12], vol. I, p. 91, where a similar result is proven for the space $\mathcal{E}$ of Schwartz.) By use of Proposition 6 we can now deduce (54).

The converse is established by the method of proof of Proposition 7.

From Proposition 8 we deduce immediately

COROLLARY. *Let* $\{G^j\}$ *be a sequence in* $\boldsymbol{Q}_A'$ *and let* M, ϵ' *be positive numbers such that, for all* j,

$$(57) \qquad |G^j(z)| \leqq b \exp(|z|^{A'-\epsilon'}) \text{ for any } z \in C.$$

Suppose also that $G^j \to 0$ *in the topology of* $\mathfrak{H}$. *Then also* $G^j \to 0$ *in the topology of* $\boldsymbol{Q}_A'$.

THEOREM 10. $\boldsymbol{Q}_A'$ *consists of all entire functions of order* $< A'$.

Proof. The fact that every function in $\boldsymbol{Q}_A'$ is an entire function of order $< A'$ was established in Proposition 5, so we proceed to the converse. Denote by K the vector space of all entire functions of order $< A'$; we have $\boldsymbol{Q}_A' \subset K$, and we want to show $\boldsymbol{Q}_A' = K$. Clearly, every polynomial lies in $\boldsymbol{Q}_A'$; we shall show that, for any $G \in K$ we can find a sequence of polynomials G^j which converges to G in the topology of $\boldsymbol{Q}_A'$. Since the space $\boldsymbol{Q}_A'$ is known to be complete, this will show that $G \in \boldsymbol{Q}_A'$.

Let us write $G(z) = \sum G_{k_1 k_2 \cdots k_n} z_1^{k_1} z_2^{k_2} \cdots z_n^{k_n}$. Then we set

$$G^j = \sum_{k_1+k_2+\cdots+k_n \leqq j} G_{k_1 k_2 \cdots k_n} Z_1^{k_1} Z_2^{k_2} \cdots Z_n^{k_n}.$$

It is clear that $G^j \to G$ in the topology of $\mathfrak{H}$. Moreover, it follows from the Corollary to Proposition 8 that $\{G^j\}$ is a Cauchy sequence in $\boldsymbol{Q}_A'$; thus, we can find a $G_1 \in \boldsymbol{Q}_A'$ such that $G^j \to G_1$ in the topology of $\boldsymbol{Q}_A'$. Now, for each $z \in C$, and each j,

$$G^j(z) = \boldsymbol{F}^{-1}(G^j) \cdot \exp(iz \cdot) \to \boldsymbol{F}^{-1}(G_1) \cdot \exp(iz \cdot) = G_1(z).$$

Thus, $G_1 = G$ which completes the proof of Theorem 10.

For any $f \in \mathfrak{Q}_A$, $z \in C$, we denote by $\tau_z f$ the function on C: $(\tau_z f)(t)$

$= f(t-z)$; it is clear that $\tau_z f \in \mathfrak{Q}_A$. For any $f \in \mathfrak{Q}_A$, $U \in \mathfrak{Q}_A'$, we define the *convolution* $U * f$ by $(U * f)(z) = U \cdot \tau_{-z} f$ for any $z \in C$. According to Proposition 4 we may write

$$(58) \qquad (U * f)(z) = \int f(t+z)\,dU(t)$$

where, for some $\epsilon > 0$, the measure U falls off like $\exp(-|t|^{A+\epsilon})$. It follows easily from (58) that $U * f \in \mathfrak{Q}_A$.

PROPOSITION 9. *If B is any bounded set in $\mathfrak{Q}_A$, then $U \to U * f$ are, for $f \in B$, equicontinuous linear maps of $\mathfrak{Q}_A' \to \mathfrak{Q}_A$. If K is any bounded set in $\mathfrak{Q}_A'$, then $f \to U * f$ are, for $U \in K$, equicontinuous linear maps of $\mathfrak{Q}_A \to \mathfrak{Q}_A$.*

Proof. By bilinearity, we need verify continuity only at zero. Let m be a fixed positive integer; since B is bounded in $\mathfrak{Q}_A$, for each integer $j \geqq m$ we can find a positive number M_j such that $v_{1/j}(f) \leqq M_j$ for all $f \in B$. Call B' the set of $f \in \mathfrak{Q}_A$ such that $v_{1/j}(f) \leqq M_j'$ for all $j \geqq m$, where $M_j' = M_j \exp(2^{A+1/m})$. Let N be the neighborhood of zero in $\mathfrak{Q}_A'$ consisting of all $U \in \mathfrak{Q}_A'$ such that $|U \cdot f| \leqq 1$ for all $f \in B'$; we claim that, for $U \in N$, $f \in B$, we have $U * f \in N_{1/m}$.

Let us note that $t \to t^{A+1/m}$ is a convex function on C. Thus, for any $t, z \in C$, any $f \in B$, and any $j > 0$, we have

$$(59) \qquad \begin{aligned} |(\tau_{-z} f)(t)| = |f(z+t)| &\leqq M_j \exp(|z+t|^{A+1/j}) \\ &\leqq M_j \exp(2^{A+1/j}(|z|^{A+1/j} + |t|^{A+1/j})). \end{aligned}$$

Thus, $\exp(-|z|^{A+1/j})(\tau_{-z} f)(t) \leqq M_j' \exp(|t|^{A+1/j})$, which implies that $\exp(-|z|^{A+1/m})\tau_{-z} f \in B'$ for all $z \in C$. Hence, for all $U \in N$, $f \in B$, $z \in C$, we have

$$|(U * f)(z)| = |U \cdot \tau_{-z} f| \leqq \exp(|z|^{A+1/m}),$$

that is, $U * f \in N_{1/m}$. This completes the proof of the first part of Proposition 9.

Let $\epsilon > 0$ be fixed. Let N' be the set of all $f \in \mathfrak{Q}_A$ such that $|U \cdot f| \leqq 1$ for all $U \in K$; by the reflexivity of $\mathfrak{Q}_A$ (Proposition 1), N' is a neighborhood of zero in $\mathfrak{Q}_A$. Thus, we can find an $\epsilon_1 > 0$ and an $a > 0$ such that $N' \subset aN_{\epsilon_1}$; call $\epsilon_2 = \min(\epsilon, \epsilon_1)$. We claim that, for $f \in \exp(-2^{A+\epsilon_2})aN_{\epsilon_2}$, $U \in K$, we have $U * f \in N_\epsilon$.

As in (59) above, if $z, t \in C$, $f \in \exp(-2^{A+\epsilon_2})aN_{\epsilon_2}$,

$$|(\tau_{-z} f)(t)| \leqq a \exp(|z|^{A+\epsilon_2} + |t|^{A+\epsilon_2})$$

so that $\exp(-|z|^{A+\epsilon_2})\tau_{-z}f \in aN_{\epsilon_2} \subset N'$; *a fortiori*, $\exp(-|z|^{A+\epsilon})\tau_{-z}f \in N'$. Thus, for any $U \in K$, $f \in \exp(-2^{A+\epsilon_2})aN_{\epsilon_2}$, we have $U * f \in N_\epsilon$ which completes the proof of Proposition 9.

For any $U, V \in \mathfrak{Q}_A'$, the convolution $U * V$ is defined by $U * V \cdot f = U \cdot V * f$ for any $f \in \mathfrak{Q}_A$. $U * V$ is in $\mathfrak{Q}_A'$ by Proposition 9.

From Proposition 9 above and the method of proof of Propositions 7 and 8 and Theorem 2 of [4] (see pp. 300-302), we can deduce

PROPOSITION 10. *$(U, V) \to U * V$ is a continuous bilinear map of $\mathfrak{Q}_A' \times \mathfrak{Q}_A' \to Q_A'$. Moreover, $U * V = V * U$.*

PROPOSITION 11. *For any $U \in \mathfrak{Q}_A'$, $f \in \mathfrak{Q}_A$, and any integers $k_1, k_2, \cdots k_n$, we have*

$$(60) \qquad (U * f)_{k_1k_2\cdots k_n} = \sum U_{l_1l_2\cdots l_n} f_{l_1+k_1, l_2+k_2, \cdots, l_n+k_n}$$

where the sum on the right is extended over all integers $l_1, l_2, \cdots, l_n$ and where the series on the right converges absolutely.

PROPOSITION 12. *For any $U, V \in \mathfrak{Q}_A'$,*

$$(61) \qquad \boldsymbol{F}(U * V) = \boldsymbol{F}(U)\boldsymbol{F}(V).$$

COROLLARY. *For $U, V, W \in \mathfrak{Q}_A'$, $U * (V * W) = (U * V) * W$, that is, convolution is associative.*

We are now in a position to study the division problem in the space $\mathfrak{Q}_A$. Since $\boldsymbol{Q}_A'$ is defined by growth conditions in C only (with no additional conditions on R) we should expect that for every $W \in \mathfrak{Q}_A'$, $W * \mathfrak{Q}_A = \mathfrak{Q}_A$. That this is actually the case will be seen later (Theorem 13 below). First, we shall characterize completely the topology of the space $\boldsymbol{Q}_A'$.

THEOREM 11. *Let H be a continuous function on C such that, for every $\epsilon' < 0$,*

$$(62) \qquad \exp(|z|^{A'-\epsilon'}) = O(H(z)).$$

Call N the set of $G \in \boldsymbol{Q}_A'$ such that $|G(z)| \leqq H(z)$ for all $z \in C$. Then these sets N form a fundamental system of neighborhoods of zero in $\boldsymbol{Q}_A'$.

Proof. Let N be a set satisfying the above hypotheses. By Proposition 7 and (62), if B is any bounded set in $\boldsymbol{Q}_A'$ we can find a positive number b such that $bB \subset N$, that is, N swallows every bounded set in $\boldsymbol{Q}_A'$. It follows from Proposition 1 and the results of [9] that $\boldsymbol{Q}_A'$ is bornologic. Thus, since N is clearly convex, N is a neighborhood of zero in $\boldsymbol{Q}_A'$.

A direct proof of the fact that N is a neighborhood of zero in $\boldsymbol{Q}_A'$ (without using the fact that Q_A' is bornologic) may also be obtained by the

same methods as those used in the proof of Theorem 14 below; this method also provides us with another proof that $\boldsymbol{Q}_A'$ is bornologic.

Conversely, let M be a neighborhood of zero in $\boldsymbol{Q}_A'$ and denote by M the set of inverse Fourier transforms of the functions in M. By the definition of the topology of $\mathfrak{Q}_A'$, we may assume that there is a bounded set K in $\mathfrak{Q}_A$ such that M consists of all $W \in \mathfrak{Q}_A'$ with $|W \cdot f| \leqq 1$ for all $f \in K$. Since K is bounded in $\mathfrak{Q}_A$, for each n-tuple of integers $(l_1, l_2, \cdots, l_n)$ we can find a positive number $c_{l_1 l_2 \cdots l_n}$ such that every $f \in K$ satisfies

$$(63) \qquad |f(z)| \leqq c_{l_1 l_2 \cdots l_n} \exp(|z_1|^{A+1/l_1} + |z_2|^{A+1/l_2} + \cdots + |z_n|^{A+1/l_n}).$$

Using Cauchy's formula we can deduce easily that every $f \in K$ satisfies

$$(64) \qquad |f_{k_1 k_2 \cdots k_n}| (k_1/(A+1/l_1))^{(k_1/(A+1/l_1))} \cdots (k_n/(A+1/k_n))^{(k_n/(A+1/l_n))}$$
$$\leqq c_{l_1 l_2 \cdots l_n} \exp[(k_1/(A+1/l_1)) + (k_2/(A+1/l_2) + \cdots$$
$$+ (k_n/(A+1/l_n))]$$

for all integers $k_1, k_2, \cdots, k_n$, $l_1, l_2, \cdots, l_n$. (We may clearly assume that $c_{00 \cdots 0} = 1$ and that $c_{l_1 l_2 \cdots l_n} \leqq c_{l_1' l_2' \cdots l_n'}$ if $l_1 \leqq l_1'$, $l_2 \leqq l_2'$, $\cdots$, $l_n \leqq l_n'$.)

Let $\{\gamma_j\}$ be a strictly increasing sequence of numbers with $\gamma_\circ = 0$; let $\{d_j\}$ be a sequence of numbers with $d_\circ = 0$, $d_j > 0$ for $j > 0$. We require that

$$\text{(a)} \qquad d_j + \gamma_j \leqq \gamma_{j+1} \text{ for any } j.$$

We define $\alpha_{j_1 j_2 \cdots j_n}$ as the region consisting of all $y \in C$ with

$$d_{j_k} + \gamma_{j_k} \leqq |y_k| \leqq \gamma_{j_k+1}$$

for all k. We require further that

$$\text{(b)} \qquad \exp(|y|^{A'-j_1'} + |y_2|^{A'-j_2'} + \cdots + |y_n|^{A'-j_n'})/c_{j_1+1, \cdots, j_n+1}$$
$$\geqq \exp(|y_1|^{A'-(j_1-1)'} + |y_2|^{A'-(j_2-1)'} + \cdots + |y_n|^{A'-(j_n-1)'})$$

whenever $|y_k| \geqq d_{j_k} + \gamma_{j_k}$ for all k and whenever $j_1 \geqq 1, \cdots, j_n \geqq 1$, where for any integer l, $A' - l'$ is defined by

$$(65) \qquad 1/(A+1/l)) + 1/(A'-l') = 1 - l^{d-1/2}$$

where $d < 0$ will be chosen later.

(c) There is no point of the form $[l/(A'-m')]^{-(A'-m')}$

in the interval $\gamma_j \leqq x \leqq \gamma_j + d_j$ *for* any j, where m, l are integers and $m \leqq 2(j+1)$.

The existences of sequences $\{\gamma_j\}$, $\{d_j\}$ satisfying (a), (b), and (c) is

not difficult to see. Moreover, it is clear that the regions $\alpha_{j_1j_2\cdots j_n}$ do not overlap.

We set

(66) $H(z)=\exp(|z_1|^{A'-j_1'}+|z_2|^{A'-j_2'}+\cdots+|z_n|^{A'-j_n'}/c_{j_1+1,j_2+1,\cdots,j_n+1}$

for z in $\alpha_{j_1j_2\cdots j_n}$. The definition of H is completed by requiring that H be continuous and monotonic, i. e. $H(z)\geqq H(z')$ if $|z_1|\geqq|z_1'|,\cdots|z_n|\geqq|z_n'|$. It is obvious from (66) and (b) above that H satisfies (62).

Call N the set of all $G\in \boldsymbol{Q}_A'$ for which $|G(z)|\leqq H(z)$ for all $z\in C$; we claim $bN\subset M$ for some $b>0$. Let $G\in N$ and let $(k_1,k_2,\cdots,k_n)$ be an n-tuple of non-negative integers. It is not difficult to see that we can find an n-tuple of integers $(l_1,l_2,\cdots,l_n)$ and an l-tuple of positive numbers $(a_1,a_2,\cdots,a_n)$ such that, for $h=1,\cdots,n$,

(67) $$[k_h/(A'-l_h')]^{-(A'-l_h')}\geqq a_h\geqq[k_h/(A'-(l_h+1)')]^{-(A'-l_h')}$$

and such that the n-chain Ξ on C defined by

$$|z_1|=a_1,\ |z_2|=a_2,\cdots,|z_n|=a_n \text{ lies in } \alpha_{l_1l_2\cdots l_n}.$$

By Cauchy's formula we have [5]

$$(\boldsymbol{F}^{-1}G)_{k_1k_2\cdots k_n}=-ik_1!k_2!\cdots k_n!\int G(z)/z_1^{k_1+1}z_2^{k_2+1}\cdots z_n^{k_n+1}\,dz.$$

Now, on Ξ we have

$$|G(z)|\leqq H(z)=\exp(|z_1|^{A'-l_1'}+|z_2|^{A'-l_2'}+\cdots+|z_n|^{A'-l_n'})/c_{l_1+1,l_2+1,\cdots,l_n+1}$$

because of (66). Using (67) we find

$$\begin{aligned}|(\boldsymbol{F}^{-1}G)_{k_1k_2\cdots k_n}|\,c_{l_1+1,l_2+1,\cdots,l_n+1}[k_1/(A'-(l_1+1)')]^{k_1(A'-l_1')}&\\ \times\cdots[k_n/(A'-(l_n+1)')]^{k_n/(A'-l_n')}&\\ \leqq k_1!k_2!\cdots k_n!\exp(k_1/(A'-l_1')+\cdots k_n/(A'-l_n')).&\end{aligned}$$

By Stirling's formula and (65) we can find a constant $\theta>0$ such that

(68) $$\begin{aligned}&(\boldsymbol{F}^{-1}G)_{k_1k_2\cdots k_n}c_{l_1+1,l_2+1,\cdots,l_n+1}\\ &\leqq\theta\exp(-k_1/(A+1/l_1)-k_2/(A+1/l_2)-\cdots-k_n/(A+1/l_n))\\ &\times k_1^{k_1/(A+1/l_1)}k_2^{k_2/(A+1/l_2)}\cdots k_n^{k_n/(A+1/l_n)}\\ &\times\exp(-k_1-k_2-\cdots-k_n)A'^{k_1/(A'-1)+k_2/(A'-1)+\cdots+k_n/(A'-1)}.\end{aligned}$$

Now, use (49), (64), (65) and (68); it is clear that we can find $d<0$ so small that, for some θ', for all $G\in N$, we have

$$|F^{-1}G\cdot f| \leqq \theta' \text{ for all } f \in K. \tag{69}$$

This proves that $(1/\theta')N \subset M$ which is the desired result.

We can now use the method of proof of Theorem 8 (except that we use Theorem 2 of [4], p. 323 in place of Theorem 5, p. 317) to deduce

THEOREM 12. *Let $J \in \boldsymbol{Q}_A'$, $J \neq 0$; then $JG \to G$ is a continuous linear map of $J\boldsymbol{Q}_A' \to \boldsymbol{Q}_A'$.*

From this we deduce easily

COROLLARY. *For any $W \in \mathfrak{Q}_A'$, $W \neq 0$, $W * U \to U$ is a continuous linear map of $W * \mathfrak{Q}_A'$ into $\mathfrak{Q}_A'$.*

THEOREM 13. *For any $W \in \mathfrak{Q}_A'$, $W * \mathfrak{Q}_A = \mathfrak{Q}_A$.*

Remark. Theorem 13 can also be derived from Theorem 4 of [4], p. 323 by use of certain *onto* theorems for Frechet spaces (see [1]). However, this method uses strongly the fact that $\mathfrak{Q}_A$ is metrizable and so does not apply to more general situations.

6. The ring of formal power series. Let $\mathfrak{O}$ denote the ring of formal power series in n variables; that is, each $f \in \mathfrak{O}$ may be written in the form $\sum f_{j_1 j_2 \cdots j_n} Z_1^{j_1} Z_2^{j_2} \cdots Z_n^{j_n}$. The topology of $\mathfrak{O}$ is that of convergence of each coefficient. It is readily verified that $\mathfrak{O}$ is a complete, Montel, bornologic, reflexive, topological vector space which is metrizable, but $\mathfrak{O}$ is not a Schwartz space. The dual $\mathfrak{O}'$ of $\mathfrak{O}$ (with the topology of uniform convergence on the bounded sets of $\mathfrak{O}$) may be identified with the space of partial differential operators; for each $U \in \mathfrak{O}'$ we write

$$U = \sum (j_1! \cdots j_n!)^{-1} U_{j_1 \cdots j_n} \partial Z_1^{j_1 + \cdots + j_n} / \partial Z_1^{j_1} \cdots \partial Z_n^{j_n}$$

(finite sum). Then we have

$$U \cdot f = \sum U_{j_1 j_2 \cdots j_n} f_{j_1 j_2 \cdots j_n}. \tag{70}$$

The convolutions $U * f$, $U * W$ for U, $W \in \mathfrak{O}'$, $f \in \mathfrak{O}$ are defined by requiring that the analog of Proposition 11 should hold for $\mathfrak{O}$; we have continuity properties similar to those in Propositions 9, 10 above.

The Fourier transform on $\mathfrak{O}'$ is defined as for the spaces $\mathfrak{H}'$, $\mathfrak{Q}_A'$. $\boldsymbol{O}' = \boldsymbol{F}\mathfrak{O}'$ may be identified with the space of polynomials in n variables. For any $U \in \mathfrak{O}'$, we have

$$\boldsymbol{F}(U) = \sum U_{j_1 j_2 \cdots j_n} Z_1^{j_1} Z_2^{j_2} \cdots Z_n^{j_n}. \tag{71}$$

The topology of $\boldsymbol{O}'$ is of interest because it is a "natural" topology for the space of polynomials.

THEOREM 14. *Let H be any continuous function on C such that, for any j, p, we have*

$$|z_j|^p = O(H(z)). \tag{72}$$

Call N the set of polynomials P such that $|P(z)| \leqq H(z)$ for all $z \in C$; the sets N form a fundamental system of neighborhoods of zero in $\boldsymbol{O}'$.

Proof. Let H and N satisfy the hypotheses of Theorem 14; we want to show that N is a neighborhood of zero in $\boldsymbol{O}'$. For each n-tuple of non-negative integers $(k_1, k_2, \cdots, k_n)$, let $a_{k_1k_2\cdots k_n}$ be any positive number such that

$$a_{k_1k_2\cdots k_n}|z_1|^{k_1}|z_2|^{k_2}\cdots|z_n|^{k_n} \leqq 2^{-k_1-k_2-\cdots-k_n}H(z)2^{-n} \tag{73}$$

for all $z \in C$. Let B be the set of all $f \in \mathfrak{O}$ for which

$$|f_{k_1k_2\cdots k_n}| \leqq a_{k_1k_2\cdots k_n} \tag{74}$$

for all $k_1, k_2, \cdots, k_n$. Call M the set of $U \in \mathfrak{O}'$ for which $|U \cdot f| \leqq 1$ for all $f \in B$; we claim that $U \in M$ implies $\boldsymbol{F}(U) \in N$.

For, given any $k_1, k_2, \cdots, k_n$, we have

$$Z_1^{k_1}Z_2^{k_2}\cdots Z_n^{k_n} \in (a_{k_1k_2\cdots k_n})^{-1}B\text{; so that } |U_{k_1k_2\cdots k_n}| \leqq a_{k_1k_2\cdots k_n}$$

or, for any $z \in C$,

$$\begin{aligned} |U(z)| &\leqq |U_{k_1k_2\cdots k_n}|\,|z_1|^{k_1}|z_2|^{k_2}\cdots|z_n|^{k_n} \\ &\leqq \sum a_{k_1k_2\cdots k_n}|z_1|^{k_1}|z_2|^{k_2}\cdots|z_n|^{k_n} \leqq \sum 2^{-k_1-k_2-\cdots-k_n}H(z)2^{-n} \leqq H(z). \end{aligned}$$

Thus, $\boldsymbol{F}(U) \in N$ which proves that N is a neighborhood of zero in $\boldsymbol{O}'$.

Call λ the topology induced by the sets N satisfying the hypotheses of Theorem 14 on the set of functions of $\boldsymbol{O}'$. Then it is clear that λ is a locally convex topology; by the above, λ is weaker than the topology of $\boldsymbol{O}'$. Now, it is readily verified that the bounded sets of λ and $\boldsymbol{O}'$ are the same, namely, the sets B of polynomials for which we can find a constant $A > 0$ such that every $P \in B$ satisfies $|P(z)| \leqq A(1 + |z|^A)$ for all $z \in C$. Thus, in order to show that the topology λ is the same as that of $\boldsymbol{O}'$, it is sufficient to show that λ defines a bornologic space. (Because it is clear that, among all locally convex topologies that can be assigned to a given vector space such that a given family of sets shall be the bounded sets, the bornologic topology is the strongest.)

Let M be any convex set in $\boldsymbol{O}'$ which swallows every bounded set in

$\boldsymbol{O}'$; we have to show that M is a neighborhood of zero for λ. For each $k_1, k_2, \cdots, k_n$, the set

$$S_{k_1k_2\cdots k_n} = \{\epsilon Z_1^{k_1}Z_2^{k_2}\cdots Z_n^{k_n}\}_{|\epsilon|\leq 1} \tag{75}$$

is bounded in $\boldsymbol{O}'$; thus, we can find a $b_{k_1k_2\cdots k_n} > 0$ with

$$b_{k_1k_2\cdots k_n} S_{k_1k_2\cdots k_n} \subset M. \tag{76}$$

Let us define the function H on C as follows: We choose a strictly increasing sequence β_k of non-negative numbers, with $\beta_0 = 0$, and a sequence of non-negative numbers d_k such that $d_0 = 0$, $d_k > 0$ for $k > 0$, $d_k + \beta_k < \beta_{k+1}$ for all k, and

$$| y_1^{k_1}y_2^{k_2}\cdots y_n^{k_n}b_{k_1k_2\cdots k_n} | \leqq | y_1^{k_1+1}y_2^{k_2+1}\cdots y_n^{k_n+1}2^{-n-k_1-k_2-\cdots-k_n} | \tag{77}$$

whenever $|y_j| \geqq d_j + \beta_j$ for all j. We call $\alpha_{k_1k_2\cdots k_n}$ the region defined by the relations

$$\alpha_{k_1k_2\cdots k_n}\colon \beta_{k_1} + d_{k_1} \leqq |z_1| \leqq \beta_{k_1+1}, \cdots, \beta_{k_n} + d_{k_n} \leqq |z_n| \leqq \beta_{k_n+1}. \tag{78}$$

Next, set

$$H(z) = 2^{-n-k_1-k_2-\cdots-k_n} |z_1|^{k_1} |z_2|^{k_2}\cdots |z_n|^{k_n} b_{k_1k_2\cdots k_n} \text{ in } \alpha_{k_1k_2\cdots k_n}. \tag{79}$$

The definition of H is completed by requiring that H be continuous and monotonic.

It is clear from (77) and (78) that H satisfies (72). Thus, the set N of $P \in \boldsymbol{O}'$ for which $|P(z)| \leqq H(z)$ is a neighborhood of zero for λ; we claim $N \subset M$. For, write $P \in N$ in the form

$$P = \sum P_{k_1k_2\cdots k_n} Z_1^{k_1}Z_2^{k_2}\cdots Z_n^{k_n}. \tag{80}$$

Given any n-tuple of non-negative integers $(l_1, l_2, \cdots, l_n)$, we can find positive numbers $a_{l_1}, a_{l_2}, \cdots, a_{l_n}$ such that the n-chain Ξ defined by: $|z_1| = a_{l_1}, |z_2| = a_{l_2}, \cdots, |z_n| = a_{l_n}$ lies in $\alpha_{l_1l_2\cdots l_n}$. By Cauchy's formula we may write [5]

$$P_{l_1l_2\cdots l_n} = i \int_{\Xi} P(z)/z_1^{l_1+1}z_2^{l_2+1}\cdots z_n^{l_n+1}\,dz.$$

Thus, since $|P(z)| \leqq H(z)$ for all z,

$$| P_{l_1l_2\cdots l_n} | \leqq H(a_{l_1}, a_{l_2}, \cdots a_{l_n})/a_{l_1}^{l_1}a_{l_2}^{l_2}\cdots a_{l_n}^{l_n} = b_{l_1l_2\cdots l_n} 2^{-n-l_1-l_2-\cdots-l_n}$$

by (79). Thus, by (75) and (76), we have

$$P_{l_1l_2\cdots l_n} Z_1^{l_1}Z_2^{l_2}\cdots Z_n^{l_n} \in 2^{-n-l_1-l_2-\cdots-l_n} M.$$

Since M is convex, we have $P \in M$; thus, $N \subset M$ which shows that λ is bornologic.

In the course of the proof of Theorem 14 we have shown

PROPOSITION 13. *$\boldsymbol{O}'$ is bornologic.*

By the method of proof of Theorem 8 (except that we use Lemma 3 of [2], p. 890 instead of Theorem 5 of [4], p. 317) we deduce

THEOREM 15. *For any polynomial P, $PG \to G$ is a continuous linear map of $P\boldsymbol{O}'$ onto $\boldsymbol{O}'$.*

We obtain easily

COROLLARY. *For each $W \in \mathcal{O}'$, $W * U \to U$ is a continuous linear map of $W * \mathcal{O}'$ onto $\mathcal{O}'$.*

THEOREM 16. *For any $W \in \mathcal{O}'$, $W * \mathcal{O}' = \mathcal{O}'$.*

Remark. Theorem 16 may also be derived easily from Theorem 1 of [4], p. 322 by use of certain onto-theorems for Frechet spaces (see [1]).

7. General remarks.

1. The methods used in Section 5 may be extended to other spaces for which the Fourier transform gives rise to entire functions. Suppose that Λ is a set and, for each $\lambda \in \Lambda$ we are given a pair of functions ϕ_λ, ψ_λ which are conjugate in the sense of Young (see [15]). Then, under suitable conditions, if we define the space $\mathfrak{Q}_\Phi$ as consisting of all entire functions f on C such that

$$f(z) = O[\exp(\phi_\lambda(z))] \text{ for all } \lambda \in \Lambda \tag{81}$$

with the topology defined by means of the semi-norms

$$v_\lambda(f) = \sup_{z \in C} |\exp(-\phi_\lambda(z))f(z)| \tag{82}$$

then we find that the space of all entire functions G on C such that

$$G(z) = O[\exp(\psi_\lambda(z))] \text{ for some } \lambda \in \Lambda \tag{83}$$

turns out to be the space $\boldsymbol{Q}_\Phi'$ which is the Fourier transform of the dual of $\mathfrak{Q}_\Phi'$. The topology of $\boldsymbol{Q}_\Phi'$ may be studied by means of the methods of Sections 4, 5, and 6, and we may also study the division problem in $\mathfrak{Q}_\Phi$.

The methods of Sections 4, 5, 6 may also be extended to other kinds of spaces, e.g. the space $\mathfrak{B}$ of functions analytic at the origin. (The topology

of $\mathfrak{B}$ is defined in a natural manner in [7].) In this case, the Fourier transform $\boldsymbol{B}'$ of the dual space is the space of all entire functions G such that

$$G(z) = O[\exp(\epsilon\,|\,z\,|)] \text{ for every } \epsilon > 0. \tag{84}$$

We can establish the analog of the theorems of Sections 4, 5, 6 for $\mathfrak{B}$.

By taking other subspaces of $\mathfrak{O}$ we obtain interesting classes of spaces of entire functions of order $\leqq 1$. These spaces are, in general, not Frechet spaces, so we cannot use the onto-theorems of [1] to prove the analog of Theorems 9, 13, and 16, but we must use the methods of proof of these theorems as given in this paper.

2. The results of Sections 2 and 3 can also be extended to other spaces. For example, the space $\mathcal{D}_\times$ which is the space of continuous linear maps of $\mathcal{D}'$ into $\mathcal{D}$ with the compact-open topology (see my paper, "On the theory of kernels of Schwartz," to appear in *Proceedings of the American Mathematical Society*). The elements of $\mathcal{D}_\times$ are the indefinitely differentiable functions of compact carrier on $R \times R$; the topology of $\mathcal{D}_\times$ is described as follows: For each integer $l \geqq 0$ we choose a finite sequence of differential operators $Q_1{}^l, Q_2{}^l, \cdots, Q^l{}_{l_j}$ on R and a number $b_l > 0$. Then we consider the set N of all $f \in \mathcal{D}_\times$ such that, if $(x,y) \in R \times R$, $|\,x\,| \geqq l_1$, $|\,y\,| \geqq l_2$,

$$|\,(Q_{k,2}{}^{l_1} Q_{k',2}{}^{l_2} f)\,(x,y)\,| \leqq b_{l_1} b_{l_2} \tag{85}$$

for $k = 1, 2, \cdots, l_{1j}$, $k' = 1, 2, \cdots, l_{2j}$, where $Q_{k,1}{}^{l_1}$ means that the operator $Q_k{}^{l_1}$ operates on x, and $Q_{k',2}{}^{l_2}$ is defined similarly for y.

The topology of the Fourier transform $\boldsymbol{D}_\times$ of $\mathcal{D}_\times$ may be described in a manner similar to that in Theorem 1, except that the regions $\Gamma_l{}^k$ which appeared there have to be replaced by regions of the form $\Gamma_l{}^k \times \Gamma_l{}^{k'}$.

3. The question remains as to what is the most general distribution of compact carrier W for which $W * \mathcal{D}' = \mathcal{D}'$, or $W * \mathcal{E} = \mathcal{E}$. We are led to a class of distributions whose Fourier transform is "slowly decreasing." This question will be dealt with in a future paper. (See "Completely inversible operators," *Proceedings of the National Academy of Science*, vol. 41 (1955), pp. 945-946.)

4. We have, in this paper, been concerned solely with complex-valued functions. It is also possible to define analogues of the spaces considered in this paper for vector-valued functions. The Fourier transform is defined as in [5]. There is no essential difficulty in describing the Fourier transform of the dual spaces in any case except for the analog of the space $\mathcal{D}$. A

description of the Fourier transform of this space in general, and even in certain particular cases seems to be extremely difficult.

5. The question naturally arises as to what extent the results of this paper can be extended to an arbitrary Lie group G in place of R or C. If G is compact, the results are very simple to obtain. The next "attackable" case is for G semi-simple. The methods of the author and F. I. Mautner (see "Some properties of the Fourier transform on semi-simple Lie groups I," *Annals of Mathematics*, vol. 61 (1955), pp. 406-439; see also the sequel to that paper which is forthcoming) together with the methods of this paper apply to certain semi-simple Lie groups. The case of the general, or solvable, Lie group seems to be very difficult.

THE INSTITUTE FOR ADVANCED STUDY.

REFERENCES.

[1] J. Dieudonné and L. Schwartz, "La dualité dans les espaces $(\mathcal{F})$ et $(\mathcal{LF})$," *Annales de l'Institut Fourier* (Grenoble), vol. I (1950), pp. 61-101.

[2] L. Ehrenpreis, "Solutions of some problems of division I," *American Journal of Mathematics*, vol. 76 (1954), pp. 883-903.

[3] ———, "Solutions of some problems of division II," *American Journal of Mathematics*, vol. 77 (1955), pp. 286-292.

[4] ———, "Mean periodic functions I," *American Journal of Mathematics*, vol. 77 (1955), pp. 293-328.

[5] ———, "Analytic functions and the Fourier transform of distributions I," *Annals of Mathematics*, vol. 63 (1956), pp. 129-159.

[6] ———, "Analytic functions and the Fourier transform of distributions II" (to appear).

[7] ———, *Theory of Infinite Derivatives* (to appear).

[8] ———, "The division problem for distributions," *Proceedings of the National Academy of Sciences*, vol. 41 (1955), pp. 756-758.

[9] A. Grothendieck, "Sur les espaces $(\mathcal{F})$ et $(\mathcal{DF})$," to appear in *Summa Brasiliensis Mathematicae*.

[10] ———, *Products Tensoriels Topologiques et Espaces Nucléaires*, Memoirs of the American Mathematical Society, No. 16, 1955.

[11] B. Malgrange, "Existence et approximation des solutions des équations aux dérivées partielles et des équations de convolution," thesis, Paris, 1955, to appear in *Annales de l'Institut Fourier* (Grenoble).

[12] L. Schwartz, *Théorie des Distributions*, vol. I-II, Paris, 1950-51.

[13] E. C. Titchmarsh, *The Theory of Functions*, Oxford, 1932.

[14] ———, *Theory of Fourier Integrals*, Oxford, 1937.

[15] A. Zygmund, *Trigonometric Series*, Warsaw, 1935.

CHARACTERISTIC LINEAR SYSTEMS OF COMPLETE CONTINUOUS SYSTEMS.*

By K. KODAIRA.†

Let V be a non-singular projective variety and let $\mathfrak{M}$ be a complete continuous system i.e., a maximal algebraic system of effective divisors D on V. Then, letting M be the parameter variety of $\mathfrak{M}$, we have a one-to-one algebraic correspondence $\lambda \rightarrow D = D_\lambda$ between M and $\mathfrak{M}$. Suppose that a member $C = D_0$ of $\mathfrak{M}$ is a non-singular prime divisor and that the corresponding point o is a simple point of M. Then the characteristic linear system of $\mathfrak{M}$ on C can be defined in a well known manner.[1] Roughly speaking. the characteristic linear system of $\mathfrak{M}$ on C is the set of all divisors $D_\lambda \cdot C$ on C cut out by divisors D_λ belonging to $\mathfrak{M}$ which are infinitely near to $C = D_0$. The main purpose of the present paper is to establish *the completeness of the characteristic linear system*[2] *of $\mathfrak{M}$ on C* under an additional restriction[3] on C. Let $\Omega(C)$ be the sheaf over V of germs of meromorphic functions which are multiples of $-C$ and denote by $H^1(V, \Omega(C))$ the first cohomology group of V with coefficients in $\Omega(C)$. Then our main results can be stated as follows:

THEOREM.[4] *Let C be a non-singular prime divisor belonging to a complete continuous system $\mathfrak{M}$ of effective divisors on a non-singular projective variety V. If C satisfies the condition*

$$H^1(V, \Omega(C)) = 0, \tag{a}$$

* Received February 14, 1956.

† This work was supported by a research project at Princeton University sponsored by the Office of Ordnance Research, U. S. Army.

[1] See Zariski [19], p. 78.

[2] For classical results concerning complete continuous systems of curves on algebraic surfaces, see Zariski [19], Chap. V.

[3] Zappa has exhibited an example of a complete continuous system of curves on an algebraic surface whose characteristic linear system on its general member is incomplete. It is therefore necessary to impose some additional restrictions on C in order to exclude such exceptional cases. See Zappa [20].

[4] In case V is an algebraic surface, this theorem reduces to a special case of a theorem of Severi to the effect that the characteristic linear system of $\mathfrak{M}$ on C is complete if $|C|$ is "semi-regular." See Severi [17].

then C corresponds to a simple point of the canonical parameter variety M of $\mathfrak{M}$ and the characteristic linear system of $\mathfrak{M}$ on C is complete. Moreover, the complete continuous system $\mathfrak{M}$ containing C is uniquely determined by C.

Let K be the canonical divisor on V. Then it can be shown that C *satisfies the condition* (a) *if*

(c′) $|m(C-K)|$ *is ample*[5] *for sufficiently large integer* m.

Hence it is a corollary of the above theorem that the characteristic linear system of the complete continuous system $\mathfrak{M}$ on its member C is complete if C belongs to the adjoint system $|K+E|$ of a system $|E|$ whose multiple $|mE|$ is ample for a sufficiently large integer[6] m.

We note that we are concerned with the *classical case* i. e. the case in which the varieties are defined over the field of complex numbers. In case the ground field is of characteristic $p \neq 0$, the theorem of the completeness of the characteristic linear systems of complete continuous systems is false, as was shown recently by J. Igusa.[7]

1. **Preliminaries.**[8] Let V be a non-singular algebraic variety of complex dimension n imbedded in a projective space. By a *complex line bundle* F over V we shall mean an analytic fibre bundle over V whose fibre is a complex line $\boldsymbol{C}$ and whose structure group is the multiplicative group $\boldsymbol{C}^*$ of complex numbers acting on $\boldsymbol{C}$. The bundle F may be described as follows: Let $\{U_j\}$ be a sufficiently fine finite covering of V and let π be the canonical projection of F onto V. Then the inverse image $\pi^{-1}(U_j)$ has a product structure: $\pi^{-1}(U_j) = U_j \times \boldsymbol{C}$, and $(z, \zeta_j) \in U_j \times \boldsymbol{C}$ is identical with $(z, \zeta_k) \in U_k \times \boldsymbol{C}$ if and only if $\zeta_j = f_{jk}(z) \cdot \zeta_k$, where $f_{jk}(z)$ is a non-vanishing holomorphic function defined in $U_j \cap U_k$. Under these circumstances, we say that the bundle F is defined by the system $\{f_{jk}\}$ of the *transition functions* $f_{jk} = f_{jk}(z)$, and we call ζ_j the *fibre coordinate* of the point (z, ζ_j) on F over the neighborhood U_j. We identify two complex line bundles which are

---

[5] We say that a complete linear system $|E|$ on V is ample if $|E|$ coincides with the system of section of V cut out by hyperplanes in one of its ambient projective spaces.

[6] This special case of our main theorem has been announced by the author at the International Congress of Mathematicians 1954. Under the stronger assumption that $|C|$ is sufficiently ample, the author has proved the completeness of $\mathfrak{M}$ on C by means of the theory of harmonic integrals. See Kodaira [8], § 11.

[7] Igusa [7a].

[8] In this Section we give a brief summary of some known results concerning complex line bundles and analytic sheaves. Cf. Kodaira [9], [10], Kodaira and Spencer [12], [13].

analytically equivalent. For any pair of bundles E, F determined respectively by $\{e_{jk}\}$, $\{f_{jk}\}$, we define the sum $E+F$ to be the bundle determined by $\{e_{jk}\cdot f_{jk}\}$. Then *the set* $\mathfrak{F}=\{F\}$ *of all complex line bundles over* V *forms an additive group.*

For any divisor D on V defined in each U_j by a local meromorphic function $R_j(D)=R_j(z;D)$, we denote by $[D]$ the line bundle over V defined by the system $\{f_{jk}(D)\}$ of the functions $f_{jk}(D)=R_j(D)/R_k(D)$. It is obvious that $[D]$ coincides with $[D']$ if and only if D is linearly equivalent to D'. Given a bundle $F\in\mathfrak{F}$, *the set of all effective divisors* D *satisfying* $[D]=F$ *forms therefore a complete linear system on* V. We denote this complete linear system by the symbol $|F|$. Obviously $|[D]|$ *coincides with* $|D|$. In view of this, we write moreover $|D+F|$ for $|[D]+F|$.

We denote by $\Gamma(F)$ the linear space consisting of all holomorphic sections of F over V. A holomorphic section $\phi\in\Gamma(F)$ is, by definition, a holomorphic mapping $z\to\phi(z)$ of V into F satisfying $\pi\phi(z)=z$. Letting $\phi_j(z)$ be the fibre coordinate of $\phi(z)$ over U_j, we infer from $\phi_j(z)=f_{jk}(z)\phi_k(z)$ that the divisor (ϕ_j) of the holomorphic function $\phi_j(z)$ coincides with the divisor (ϕ_k) of $\phi_k(z)$ in $U_j\cap U_k$. Hence we may define *the divisor* (ϕ) *of the section* ϕ by setting $(\phi)=(\phi_j)$ on each U_j, provided that $\phi\neq 0$. As one readily infers, the complete linear system $|F|$ consists of all divisors $D=(\phi)$, $\phi\in\Gamma(F)$:

$$|F|=\{(\phi)\,|\,\phi\in\Gamma(F),\neq 0\}. \tag{1.1}$$

Obviously (ψ) coincides with (ϕ) if and only if $\psi=c\cdot\phi$, $c\in\boldsymbol{C}^*$. Consequently we obtain

$$\dim|F|=\dim\Gamma(F)-1. \tag{1.2}$$

By the *characteristic class* $c(F)$ of a complex line bundle F over V we shall mean the characteristic class of the principal bundle associated with F. The characteristic class $c(F)$ is an element of the second cohomology group $H^2(V,\boldsymbol{Z})$ of V with coefficients in the integers $\boldsymbol{Z}$. Let $\boldsymbol{R}$ be the reals. For any element $c\in H^2(V,\boldsymbol{Z})$ we denote by c_R the element of $H^2(V,\boldsymbol{R})$ corresponding to c under the homomorphism $H^2(V,\boldsymbol{Z})\to H^2(V,\boldsymbol{R})$ induced by the inclusion map $\boldsymbol{Z}\to\boldsymbol{R}$. In view of de Rham's theorem, c_R can be regarded as a class of closed real 2-forms on V.

THEOREM 1.1.[9] *Let* F *be the complex line bundle over* V *defined by a system* $\{f_{jk}\}$ *of transition functions* f_{jk} *with respect to a covering* $\{U_j\}$,

[9] Kodaira [10], p. 1271, Lemma.

and let $\{a_j\}$ *be a system of positive functions* a_j *defined respectively on* U_j *satisfying* $a_j = |f_{jk}|^2 a_k$ *in* $U_j \cap U_k$. *Then* $\gamma = (\frac{i}{2\pi})\partial\bar{\partial} \log a_j$ *is a closed real* $(1,1)$*-form on* V *belonging to the characteristic class* $c_R(F)$ *of the bundle* F.

We say that a real $(1,1)$-form $\gamma = i\sum \gamma_{\alpha\bar{\beta}}(z,\bar{z})\,dz_\alpha d\bar{z}_\beta$ on V is *positive* and write $\gamma > 0$ if the Hermitian form $\sum \gamma_{\alpha\bar{\beta}}(z,\bar{z})u_\alpha \bar{u}_\beta$ in n variables $u_1, \cdots, u_n$ is positive definite at each point z on V. Moreover, we say that a cohomology class $c \in H^2(V, \boldsymbol{Z})$ is positive and write $c > 0$ if c_R contains a closed real $(1,1)$-form $\gamma > 0$. By the canonical bundle K over V we shall mean the complex line bundle of $(n,0)$-forms over V, where n is the dimension of V. Now we denote by $\Omega(F)$ the sheaf of germs of holomorphic sections of the bundle F. Then we have

THEOREM 1.2.[10] *The cohomology group* $H^q(V, \Omega(F))$ *vanishes for* $q \geqq 1$ *if the characteristic class* $c(F-K)$ *of* $F-K$ *is positive.*

We say that a complete linear system $|D|$ on V is *ample* if there exists an ambient projective space $\mathfrak{S}$ of V such that $|D|$ coincides with the system of all sections of V cut out by the hyperplanes in $\mathfrak{S}$.

THEOREM 1.3.[11] *There exists on* V *a real* $(1,1)$*-form* γ_0 *such that the complete linear system* $|F|$ *is ample for any complex line bundle* F *whose characteristic class* $c_R(F)$ *contains a closed real* $(1,1)$*-form* $\gamma > \gamma_0$.

Combining this with the obvious fact that $c(F)$ is positive if $|F|$ is ample, we infer that the *characteristic class* $c(F)$ *of* F *is positive if and only if the complete linear system* $|mF|$ *is ample for sufficiently large positive integer* m.

Let $\mathfrak{A}$ be the linear space over $\boldsymbol{C}$ consisting of all simple differentials of the first kind on V. The first cohomology group $H^1(V, \boldsymbol{Z})$ is free abelian. Let $\{b_1, \cdots, b_r, \cdots, b_{2q}\}$ be a base of $H^1(V, \boldsymbol{Z})$. Then, for each b_r, there exists one and only one element β_r of $\mathfrak{A}$ such that

$$2\pi i b_r[Z] = \int_Z (\bar{\beta}_r - \beta_r) \tag{1.3}$$

for any integral 1-cycle Z on V. Clearly $\{\beta_1, \cdots, \beta_r, \cdots, \beta_{2q}\}$ forms a base of $\mathfrak{A}$ with respect to the reals $\mathbf{R}$. We denote by $\mathfrak{d}$ the discrete subgroup of $\mathfrak{A}$ generated by $\beta_1, \cdots, \beta_r, \cdots, \beta_{2q}$. Now let $\mathfrak{P}$ be the subgroup of $\mathfrak{F}$

[10] Kodaira [11], Theorem 3.

[11] Kodaira [11], Theorem 3.

consisting of all complex lines bundles P with $c(P)=0$. Then we have the exact sequence[12]

$$(1.4)\qquad 0 \to \bar{\mathfrak{D}} \xrightarrow{i} \bar{\mathfrak{A}} \xrightarrow{\rho} \mathfrak{P} \to 0,$$

where i denotes the inclusion map. The homomorphism ρ is described explicitly as follows: For $\bar{\alpha} \in \bar{\mathfrak{A}}$, the image $P=\rho(\bar{\alpha})$ is the complex line bundle over V defined with respect to the covering $\{U_j\}$ by the system $\{\kappa_{jk}\}$ of the constants

$$(1.5)\qquad \kappa_{jk} = \exp \int_{z(j)}^{z(k)} \bar{\alpha},$$

where $z(j)$ is a point in U_j and $\int_{z(j)}^{z(k)}$ denotes the integral along a smooth curve *in* $U_j \cup U_k$ combining $z(j)$ with $z(k)$. The canonical complex structure on $\mathfrak{P}$ is obtained by identifying $\mathfrak{P}$ with the complex torus $\bar{\mathfrak{A}}/\bar{\mathfrak{D}}$ by means of the isomorphism $\bar{\mathfrak{A}}/\bar{\mathfrak{D}} \cong \mathfrak{P}$ induced by ρ. The complex manifold $\mathfrak{P}=\bar{\mathfrak{A}}/\bar{\mathfrak{D}}$ is called the *Picard variety* attached to V. It is well known that the *Picard variety* $\mathfrak{P}$ *is an algebraic variety imbedded in a projective space.*[13]

2. Construction of a complex line bundle. We form the product variety $V \times \mathfrak{P}$ and denote for any point $P \in \mathfrak{P}$ the mapping $z \to (z, P)$ of V into $V \times \mathfrak{P}$ by r_P. In this Section we construct a complex line bundle Ξ over $V \times \mathfrak{P}$ such that, for each complex line bundle $P \in \mathfrak{P}$, the bundle $r_P^*(\Xi)$ over V induced by the mapping r_P coincides with P: $r_P^*(\Xi) = P$.

Let M be a compact complex manifold, $\hat{M}$ the universal covering of M, and let Δ be the covering transformation group of $\hat{M}$ with respect to M. Moreover, let $f(\hat{z}, \tau)$, $\hat{z} \in \hat{M}$, $\tau \in \Delta$, be a non-vanishing complex-valued function which is holomorphic in $\hat{z}$ and satisfies

$$(2.1)\qquad f(\hat{z}, \sigma\tau) = f(\tau\hat{z}, \sigma) \cdot f(\hat{z}, \tau), \qquad \sigma, \tau \in \Delta.$$

Then, defining for each $\tau \in \Delta$ an analytic automorphism $T(\tau)$ of the product variety $\hat{M} \times \boldsymbol{C}$ by

$$(2.2)\qquad T(\tau):\ (\hat{z}, \zeta) \to (\tau\hat{z}, f(\hat{z}, \tau)\zeta),$$

we infer readily that $T(\sigma\tau^{-1}) = T(\sigma)T(\tau)^{-1}$ for $\sigma, \tau \in \Delta$. Thus

$$T(\Delta) = \{T(\tau) \mid \tau \in \Delta\}$$

---

[12] Kodaira and Spencer [12], p. 871.

[13] Lefschetz [15], pp. 364-370; see also Kodaira [11], p. 40.

forms a discontinuous group of analytic automorphisms of $\hat{M} \times \boldsymbol{C}$ which is isomorphic to Δ, and the factor space $F = \hat{M} \times \boldsymbol{C}/T(\Delta)$ is a complex manifold (without singularities). Obviously each automorphism $T(\tau)$ commutes with the canonical projection $\hat{\pi}: (\hat{z}, \zeta) \rightarrow \hat{z}$ of $\hat{M} \times \boldsymbol{C}$ onto $\hat{M}$. Hence $\hat{\pi}$ induces a holomorphic mapping π of F onto $M = \hat{M}/\Delta$, and moreover, for a sufficiently small open subset $U \subset M$, the inverse image $\pi^{-1}(U)$ has a canonical product structure $\pi^{-1}(U) = U \times \boldsymbol{C}$. Thus we infer that F is a complex line bundle[14] over M.

We remark that a system $\{f_{jk}\}$ of transition functions defining the bundle F can be obtained in the following manner: Denote by ϖ the canonical projection of $\hat{M}$ onto M. Letting $\{U_j\}$ be a finite covering of M by sufficiently small "spherical" neighborhoods U_j, we take, for each U_j, a spherical neighborhood $\hat{U}_j$ on $\hat{M}$ which is mapped by ϖ homeomorphically onto U_j, and denote by ϖ_j the restriction of ϖ to $\hat{U}_j$. Then, for each pair (j, k) such that $U_j \cap U_k$ is not empty, the mapping $\varpi_j^{-1}\varpi_k$ defined on $\hat{U}_k \cap \varpi^{-1}(U_j)$ coincides with an element τ_{jk} of Δ:

$$\varpi_j^{-1}\varpi_k = \tau_{jk}, \qquad \text{on } \hat{U}_k \cap \varpi^{-1}(U_j).$$

Now, setting

$$f_{jk}(z) = f(\varpi_k^{-1}(z), \tau_{jk}), \qquad \text{for } z \in U_j \cap U_k, \tag{2.3}$$

we infer readily that $\{f_{jk}(z)\}$ is a system of transition functions defining the complex line bundle $F = \hat{M} \times \boldsymbol{C}/T(\Delta)$. In fact, denote by $[\hat{z}, \zeta]$ the point on F corresponding to $(\hat{z}, \zeta) \in \hat{M} \times \boldsymbol{C}$. Then each point η on the fibre $\pi^{-1}(z)$, $z \in U_j$, can be written uniquely in the form $\eta = [\varpi_j^{-1}(z), \zeta_j]$. Thus we may use ζ_j as the fibre coordinate of η over U_j. Now, in view of (2.2), we have

$$[\varpi_k^{-1}(z), \zeta_k] = [\tau_{jk}\varpi_k^{-1}(z), f(\varpi_k^{-1}(z), \tau_{jk})\zeta_k]$$

and therefore, by (2.3),

$$[\varpi_k^{-1}(z), \zeta_k] = [\varpi_j^{-1}(z), f_{jk}(z)\zeta_k] \qquad \text{for } z \in U_j \cap U_k.$$

This proves that the law of transition of fibre coordinates is given by $\zeta_j = f_{jk}(z)\zeta_k$.

LEMMA 2.1. *Let $a(\hat{z})$ be a positive function of class C^∞ on $\hat{M}$ satisfying*

$$a(\tau\hat{z}) = |f(\hat{z}, \tau)|^2 \cdot a(\hat{z}), \quad \tau \in \Delta. \tag{2.4}$$

---

[14] For the general theory of invariant complex line bundles over complex manifolds with discontinuous groups of analytic automorphisms, see Baily [1], [2].

Then the closed real $(1,1)$*-form* $\gamma=(\frac{i}{2\pi})\partial\bar{\partial}\log a(\hat{z})$ *is invariant under* Δ. *Moreover, considered as a closed real* $(1,1)$*-form on* $M=\hat{M}/\Delta$, γ *belongs to the characteristic class* $c_R(F)$ *of the bundle* F.

Proof. It is obvious that γ is invariant under Δ. Now setting

$$a_j(z)=a(\varpi_j^{-1}(z)), \qquad z\in U_j,$$

we infer from (2.3) and (2.4) that

$$a_j(z)=|f_{jk}(z)|^2 a_k(z), \qquad z\in U_j\cap U_k,$$

while we have $\gamma=(\frac{i}{2\pi})\partial\bar{\partial}\log a_j(z)$. Hence, by Theorem 1.1, γ belongs to $c_R(F)$, q. e. d.

In order to construct the complex line bundle Ξ, we apply the above procedure to the product manifold $M=V\times\mathfrak{P}$. Denote by $\hat{V}$ the universal covering of V, by G the covering transformation group of $\hat{V}$ with respect to V, and by ϖ the canonical projection of $\hat{V}$ onto V. Then, since $\mathfrak{P}=\bar{\mathfrak{A}}/\bar{\mathfrak{d}}$, the universal covering $\hat{M}$ of $M=V\times\mathfrak{P}$ is given by $\hat{M}=\hat{V}\times\bar{\mathfrak{A}}$ and the covering transformation group Δ of $\hat{M}$ is given by $\Delta=G\times\bar{\mathfrak{d}}$. Letting $\hat{o}$ be a point on $\hat{V}$ fixed once and for all and denoting by $C(\hat{o},\hat{z})$ an oriented continuous curve combining $\hat{o}$ with $\hat{z}$, where $\hat{z}$ is a point on V, we associate with each element $g\in G$ the closed continuous curve $C(g)=\varpi C(\hat{o},g\hat{o})$ on V. As is well known, $g\to C(g)$ induces an isomorphism of G onto the fundamental group of V. Now we set

$$f(\hat{z},\bar{\alpha};g,\bar{\delta})=\exp\{\int_{C(g)}(\bar{\alpha}+\bar{\delta})+\int_{\hat{o}}^{\hat{z}}\varpi^*\delta\},$$

where $(\hat{z},\bar{\alpha})\in\hat{V}\times\bar{\mathfrak{A}}=\hat{M}$, $(g,\bar{\delta})\in G\times\bar{\mathfrak{d}}=\Delta$ and where $\varpi^*\delta$ is the holomorphic 1-form on $\hat{V}$ induced by $\delta=\bar{\bar{\delta}}$. The non-vanishing function $f(\hat{z},\bar{\alpha};g,\bar{\delta})$ is obviously holomorphic in $(\hat{z},\bar{\alpha})$. Moreover $f(\hat{z},\bar{\alpha};g,\bar{\delta})$ satisfies the relation (2.1). In fact, we have

$$\begin{aligned} & f(h\hat{z},\bar{\alpha}+\bar{\epsilon};g,\bar{\delta})\cdot f(\hat{z},\bar{\alpha};h,\bar{\epsilon}) \\ &=\exp\{\int_{C(g)}(\bar{\alpha}+\bar{\epsilon}+\bar{\delta})+\int_{\hat{o}}^{h\hat{z}}\varpi^*\delta+\int_{C(h)}(\bar{\alpha}+\bar{\epsilon})+\int_{\hat{o}}^{\hat{z}}\varpi^*\epsilon\} \\ &=\exp\{\int_{C(gh)}(\bar{\alpha}+\bar{\delta}+\bar{\epsilon})+\int_{\hat{o}}^{\hat{z}}\varpi^*(\delta+\epsilon)-\int_{C(h)}\bar{\delta}+\int_{\hat{z}}^{h\hat{z}}\varpi^*\delta\}, \end{aligned}$$

where $h\in G$, $\bar{\epsilon}\in\bar{\mathfrak{d}}$, while, since δ has the form $\delta=\sum m_r\beta_r$, $m_r\in\mathbf{Z}$, we get, using (1.3),

$$\int_{C(h)} \bar{\delta} - \int_{\hat{z}}^{h\hat{z}} \varpi^* \delta = \int_{C(h)} (\bar{\delta} - \delta) = \sum m_r \int_{C(h)} (\bar{\beta}_r - \beta_r) = 2\pi i \sum m_r b_r [C(h)],$$

where $\sum m_r b_r [C(h)]$ is a rational integer. Hence we obtain

$$f(h\hat{z}, \bar{\alpha} + \bar{\epsilon}; g, \bar{\delta}) \cdot f(\hat{z}, \bar{\alpha}; h, \bar{\epsilon}) = f(\hat{z}, \bar{\alpha}; gh, \bar{\delta} + \bar{\epsilon}).$$

Letting $T(g, \delta)$ be the analytic automorphism of $\hat{M} \times \boldsymbol{C} = \hat{V} \times \bar{\mathfrak{A}} \times \boldsymbol{C}$ defined by

$$T(g, \bar{\delta}) : (\hat{z}, \bar{\alpha}, \zeta) \to (g\hat{z}, \bar{\alpha} + \bar{\delta}, f(\hat{z}, \bar{\alpha}; g, \bar{\delta}) \cdot \zeta),$$

we infer therefore that $T(G \times \bar{\mathfrak{d}}) = \{T(g, \bar{\delta}) \mid g \in G, \bar{\delta} \in \bar{\mathfrak{d}}\}$ is a discontinuous group of analytic automorphisms of $\hat{V} \times \bar{\mathfrak{A}} \times \boldsymbol{C}$ and that the factor space

$$\Xi = \hat{V} \times \bar{\mathfrak{A}} \times \boldsymbol{C} / T(G \times \bar{\mathfrak{d}})$$

is a complex line bundle over $V \times \mathfrak{P} = \hat{V} \times \bar{\mathfrak{A}} / G \times \bar{\mathfrak{d}}$.

In order to show that the induced bundle $r_P^*(\Xi)$ over V coincides with P, we take an element $\bar{\alpha} \in \bar{\mathfrak{A}}$ such that $\rho(\bar{\alpha}) = P$. Then we have

$$r_P^*(\Xi) = \hat{V} \times \bar{\alpha} \times \boldsymbol{C} / T(G \times \bar{0}).$$

By the canonical identification $\hat{V} \times \bar{\alpha} \times \boldsymbol{C} = \hat{V} \times \boldsymbol{C}$, each automorphism $T(g, \bar{0})$ is reduced to

$$T(g) : (\hat{z}, \zeta) \to (g\hat{z}, f(\hat{z}, g) \cdot \zeta),$$

where $f(\hat{z}, g) = f(\hat{z}, \bar{\alpha}; g, \bar{0}) = \exp \int_{C(g)} \bar{\alpha}$. Thus the bundle $r_P^*(\Xi)$ is represented in the form $r_P^*(\Xi) = \hat{V} \times \boldsymbol{C} / T(G)$. Now take a finite covering $\{U_j\}$ of V by sufficiently small spherical neighborhoods U_j. To find a system $\{f_{jk}(z)\}$ of transition functions defining the bundle $r_P^*(\Xi)$ with respect to $\{U_j\}$, we apply the formula (2.3) to $r_P^*(\Xi)$, Then, letting ϖ_j have the same meaning as in (2.3), we get

$$f_{jk}(z) = \exp \int_{C(g_{jk})} \bar{\alpha}, \qquad z \in U_j \cap U_k,$$

where $g_{jk} = \varpi_j^{-1} \varpi_k$ on $\varpi_k^{-1}(U_k)$. Take a point $z(j)$ in each U_j and set $\hat{z}(j) = \varpi_j^{-1}(z(j))$. Then we have

$$\int_{C(g_{jk})} \bar{\alpha} = \int_{\varpi_k^{-1}(z)}^{\varpi_j^{-1}(z)} \varpi^* \bar{\alpha} = \left\{ \int_{\varpi_k^{-1}(z)}^{\hat{z}(k)} + \int_{\hat{z}(k)}^{\hat{z}(j)} + \int_{\hat{z}(j)}^{\varpi_j^{-1}(z)} \right\} \varpi^* \bar{\alpha}.$$

where N_P^*, Π^*, i^* denote the homomorphisms induced by N_P, Π, i, respectively. In this diagram (3.7), N_P^* and Π^* are both *isomorphisms onto.* In fact, it is obvious by (3.6) that Π^* maps $\Gamma(E+P)$ isomorphically onto $\Gamma(\tilde{F}_{V\times S})$. N_P^* is clearly an isomorphism of $\Gamma(F)$ *into* $\Gamma(\tilde{F})$. Take an arbitrary section $\psi\in\Gamma(\tilde{F})$. Then, since N_P is a biregular map of $V\times\tilde{\mathfrak{P}}-V\times S$ onto $V\times\mathfrak{P}-V\times P$, there exists a holomorphic section ϕ' of F over $V\times\mathfrak{P}-V\times P$ such that $N_P^*\phi'=\psi$ on $V\times\tilde{\mathfrak{P}}-V\times S$. By virtue of a theorem of Hartogs, ϕ' can be extended to $\phi\in\Gamma(F)$. This extension ϕ satisfies obviously $N_P^*\phi=\psi$ on $V\times\tilde{\mathfrak{P}}$. Thus we see that N_P^* maps $\Gamma(F)$ isomorhpically *onto* $\Gamma(\tilde{F})$. To prove our lemma, it is therefore sufficient to show that

$$(3.8)\qquad i^*\Gamma(\tilde{F})=\Gamma(\tilde{F}_{V\times S}).$$

We observe the exact sequence

$$0\to\Omega(\tilde{F}-[V\times S])\to\Omega(\tilde{F})\xrightarrow{r_{V\times S}}\Omega(\tilde{F}_{V\times S})\to 0.$$

The exact cohomology sequence corresponding to this can be written in the form

$$\cdots\to\Gamma(\tilde{F})\xrightarrow{i^*}\Gamma(\tilde{F}_{V\times S})\to H^1(V\times\tilde{\mathfrak{P}},\Omega(\tilde{F}-[V\times S]))\to\cdots$$

since the homomorphism $r^*_{V\times S}$: $\Gamma(\tilde{F})\to\Gamma(\tilde{F}_{V\times S})$ coincides with i^*. In order to show (3.8), it is therefore sufficient to prove that

$$(3.9)\qquad H^1(V\times\tilde{\mathfrak{P}},\Omega(\tilde{F}-[V\times S]))=0.$$

For simplicity's sake we use the following abbreviations: Let M, N be complex manifolds. Given a complex line bundle B [or a form γ] on M, we denote by the same symbol B [or γ] simultaneously the bundle [or the form] on the product space $M\times N$ induced by B [or γ] in an obvious manner. For example we mean by the bundle K over $V\times\mathfrak{P}$ the bundle over $V\times\mathfrak{P}$ induced by the canonical bundle K over V. Now we prove (3.9) with the help of Theorem 1.2. Since the canonical bundle over $\mathfrak{P}$ is trivial, the canonical bundle over $\tilde{\mathfrak{P}}$ is equal[21] to $(q-1)[S]$. Hence the *canonical bundle* $\mathfrak{K}$ *over* $V\times\tilde{\mathfrak{P}}$ is given by $\mathfrak{K}=K+(q-1)[V\times S]$. We have therefore

$$(3.10)\qquad \tilde{F}-[V\times S]-\mathfrak{K}=E-K+\tilde{\Xi}+\tilde{\mathfrak{E}}-q[V\times S],$$

where $\tilde{\Xi}=N_P^*(\Xi)$, $\tilde{\mathfrak{E}}=N_P^*(\mathfrak{E})$. Setting $m=m'+m''$, we decompose $\mathfrak{E}$ in the right hand side of (3.10) into two parts:

$$\mathfrak{E}=\mathfrak{E}'+\mathfrak{E}'',\qquad\qquad \text{where } \mathfrak{E}'=m'\mathfrak{E}_1,\ \mathfrak{E}''=m''\mathfrak{E}_1.$$

---

[21] Kodaira [11], p. 31.

By hypothesis $c_R(E-K)$ contains a closed *positive* $(1,1)$-form γ on V. We have therefore

$$c_R(E-K+\Xi+\mathfrak{E}')\ni\gamma+\xi+m'\omega, \qquad \text{on } V\times\mathfrak{P},$$

where ξ is the form defined by (2.6) belonging to $c_R(\Xi)$. Clearly the restriction $\xi_{V\times Q}$ of ξ to $V\times Q$ vanishes for each point $Q\in\mathfrak{P}$, while ω is positive on $\mathfrak{P}$. Hence we get

$$\gamma+\xi+m'\omega>0. \qquad \text{on } V\times\mathfrak{P},$$

provided that m' is sufficiently large. On the other hand, $c_R(\tilde{\mathfrak{E}}''-q[S])$ contains a closed positive $(1,1)$-form[22] ρ'' on $\tilde{\mathfrak{P}}$, provided that m'' is sufficiently large, where $\tilde{\mathfrak{E}}''=\sigma_P(\mathfrak{E}'')$ is the bundle over $\tilde{\mathfrak{P}}$ induced by $\mathfrak{E}''=m''\mathfrak{E}_1$. Setting $\rho=\gamma+\xi+m'\omega$ on $V\times\mathfrak{P}$, we therefore infer from (3.10) that

$$c_R(\tilde{F}-[V\times S]-\mathfrak{R})\ni\tilde{\rho}+\rho'', \qquad \text{on } V\times\tilde{\mathfrak{P}},$$

where $\tilde{\rho}=N_P{}^*(\rho)$. To show that $\tilde{\rho}+\rho''$ is positive on $V\times\tilde{\mathfrak{P}}$, we observe that

$$\Psi:\ (z,\tilde{Q})\to(z,Q,\tilde{Q}), \qquad Q=\sigma_P{}^{-1}(\tilde{Q}),$$

is a bi-regular map of $V\times\tilde{\mathfrak{P}}$ into $V\times\mathfrak{P}\times\tilde{\mathfrak{P}}$, where $(z,\tilde{Q})$ denotes a "variable" point on $V\times\tilde{\mathfrak{P}}$. Since ρ or ρ'' is positive on $V\times\mathfrak{P}$ or $\tilde{\mathfrak{P}}$, the sum $\rho+\rho''$ is positive on $V\times\mathfrak{P}\times\tilde{\mathfrak{P}}$, while $\tilde{\rho}+\rho''=\Psi^*(\rho+\rho'')$. Consequently $\tilde{\rho}+\rho''$ is positive on $V\times\tilde{\mathfrak{P}}$. Thus we see that

$$c_R(\tilde{F}-[V\times S]-\mathfrak{R})>0, \qquad \text{on } V\times\tilde{\mathfrak{P}}.$$

and therefore, by using Theorem 1.2, we obtain (3.9), q.e.d.

LEMMA 3.2.[23] *Let B be an arbitrary complex line bundle over V and let Q be a member of $\mathfrak{P}$. Then we have*

$$\text{(3.11)} \qquad \limsup\dim\Gamma(B+P)\leqq\dim\Gamma(B+Q)\ \textit{as}\ P\to Q,$$

where P is a "variable" member of $\mathfrak{P}$.

Proof. Take a sufficiently fine covering $\{U_j\}$ of V. Then the complex line bundle P is defined by the system $\{\kappa_{jk}\}$ of the constants

$$\kappa_{jk}=\exp\int_{z(j)}^{z(k)}\tilde{\alpha},$$

[22] Kodaira [11], p. 31, Lemma 1.

[23] This lemma is reduced to a special case of a theorem of Kodaira and Spencer concerning semi-continuity of the dimensions of cohomologies. See Kodaira and Spencer [14], p. 15. We give here a simplified version of their proof for this special case.

where $\bar{\alpha}$ is an element of $\bar{\mathfrak{A}}$ such that $\rho(\bar{\alpha}) = P$ (see (1.5)). The system $\{\bar{\kappa}_{jk}\}$ defines an analytically trivial bundle, since

$$\bar{\kappa}_{jk} = f_j(z)/f_k(z), \qquad \text{where } f_j(z) = \exp \int_{z(j)}^{z} \alpha.$$

Hence we may replace $\{\kappa_{jk}\}$ by the system $\{\theta_{jk}\}$ of constants $\theta_{jk} = \kappa_{jk}/\bar{\kappa}_{jk}$ with $|\theta_{jk}| = 1$. Now let $\{b_{jk}(z)\}$ be the system of transition functions defining the bundle B and let $\{a_j(z)\}$ be the system of functions $a_j(z) > 0$ of class C^∞ on U_j such that $a_j(z) = |b_{jk}(z)|^2 a_k(z)$ in $U_j \cap U_k$. For any section $\phi: z \to \phi(z)$ belonging to $\Gamma(B+P)$, the fibre coordinates $\phi_j(z)$ of $\phi(z)$ satisfy

$$\phi_j(z) = b_{jk}(z)\theta_{jk} \cdot \phi_k(z), \qquad \text{in } U_j \cap U_k,$$

and therefore $a_j(z)^{-1} |\phi_j(z)|^2 = a_k(z)^{-1} |\phi_k(z)|^2$. Hence we may define the inner product

$$(\phi', \phi) = \int_V a_j(z)^{-1} \phi_j'(z) \bar{\phi}_j(z)\, dV, \qquad \text{for } \phi', \phi \in \Gamma(B+P),$$

where dV denotes the invariant volume element on V. Now let

$$l = \limsup \dim \Gamma(B+P) \text{ as } P \to Q.$$

Then there exists a sequence $P_1, P_2, \cdots, P_m, \cdots$ with $\lim P_m = Q$ such that $\dim \Gamma(B + P_m) = l$ for $m = 1, 2, \cdots$. We take $\bar{\alpha}_m, \bar{\alpha}_\infty \in \bar{\mathfrak{A}}$ with $\rho(\bar{\alpha}_m) = P_m$, $\rho(\bar{\alpha}_\infty) = Q$ in such a way that $\bar{\alpha}_m \to \bar{\alpha}_\infty$ for $m \to \infty$, and denote by $\theta_{jk}^{(m)}$ the θ_{jk} corresponding to $\bar{\alpha}_m$. Moreover we form a base

$$\{\phi_1^{(m)}, \cdots, \phi_\nu^{(m)}, \cdots, \phi_l^{(m)}\}$$

of each $\Gamma(B + P_m)$ satisfying $(\phi_\lambda^{(m)}, \phi_\nu^{(m)}) = \delta_{\lambda\nu}$. We have

$$\int_{U_j} a_j(z)^{-1} |\phi_{\nu j}^{(m)}(z)|^2\, dV \leqq (\phi_\nu^{(m)}, \phi_\nu^{(m)}) = 1.$$

There exists therefore a subsequence [24] of $\{\phi_{\nu j}^{(1)}, \phi_{\nu j}^{(2)}, \cdots, \phi_{\nu j}^{(m)}, \cdots\}$ which converges uniformly on each compact subset of U_j. Consequently, by a suitable choice of the sequence $\{P_m\}$, we may assume that, as $m \to \infty$, the limit $\phi_{\nu j}(z) = \lim \phi_{\nu j}^{(m)}(z)$ exists and represents a holomorphic function on U_j. Since

$$\phi_{\nu j}^{(m)}(z) = b_{jk}(z)\theta_{jk}^{(m)} \cdot \phi_{\nu k}^{(m)}(z), \qquad \text{in } U_j \cap U_k,$$

we have

$$\phi_{\nu j}(z) = b_{jk}(z)\theta_{jk}^{(\infty)} \cdot \phi_{\nu k}(z), \qquad \text{in } U_j \cap U_k,$$

[24] Bochner and Martin [3], pp. 116-118.

where $\theta_{jk}^{(\infty)} = \lim \theta_{jk}^{(m)}$ is the θ_{jk} corresponding to $\bar{a}_{\infty}$. This shows that $\phi_{\nu j}(z)$ represents the fibre coordinate over U_j of a section $\phi_\nu \in \Gamma(B+Q)$. Moreover we have

$$(\phi_\lambda, \phi_\nu) = \lim_{m\to\infty} (\phi_\lambda^{(m)}, \phi_\nu^{(m)}) = \delta_{\lambda\nu}.$$

Thus we infer that $\Gamma(B+Q)$ contains l linearly independent elements $\phi_1, \cdots, \phi_\nu, \cdots, \phi_l$ and therefore we obtain (3.11).

LEMMA 3.3. *The dimension of* $\Gamma(E+P)$ *is independent of* $P \in \mathfrak{P}$, *provided that* $c(E-K)$ *is positive on* V.

Proof. Take an arbitrary point $Q \in \mathfrak{P}$ and set $\dim \Gamma(E+Q) = d+1$. We readily infer from the above Lemma 3.1 that

$$\liminf \dim \Gamma(E+P) \geqq d+1 \text{ as } P \to Q. \tag{3.12}$$

In fact, since every element ϕ of $\Gamma(E+Q)$ can be written in the form $\phi = r_Q{}^*\psi$, $\psi \in \Gamma(F)$, there exist $d+1$ elements $\psi_0, \psi_1, \cdots, \psi_d$ of $\Gamma(F)$ such that $r_Q{}^*\psi_0, r_Q{}^*\psi_1, \cdots, r_Q{}^*\psi_d$ are linearly independent. It follows that $d+1$ elements $r_P{}^*\psi_0, r_P{}^*\psi_1, \cdots, r_P{}^*\psi_d$ of $\Gamma(E+P)$ are linearly independent if P lies in a sufficiently small neighborhood of Q. Hence we obtain the inequality (3.12). Now, comparing (3.12) with (3.11), we infer immediately that the integer-valued function $\dim \Gamma(E+P)$ in P reduces to a constant, q. e. d.

By means of the above Lemmas 3.1 and 3.3 we define an analytic bundle structure on the set $\mathcal{B} = \bigcup_{P\in\mathfrak{P}} \Gamma(E+P)$ in the following manner: First we define the projection π: $\mathcal{B} \to \mathfrak{P}$ by $\pi(\phi) = P$ for $\phi \in \Gamma(E+P)$. Given a point $Q \in \mathfrak{P}$, we take $d+1$ elements $\psi_0, \psi_1, \cdots, \psi_d \in \Gamma(F)$ such that $r_Q{}^*\psi_0, r_Q{}^*\psi_1, \cdots, r_Q{}^*\psi_d$ form a base of $\Gamma(E+Q)$, where $d+1 = \dim \Gamma(E+Q)$. This is possible because of Lemma 3.1. Obviously $r_P{}^*\psi_0, \cdots, r_P{}^*\psi_d \in \Gamma(E+P)$ are linearly independent if P lies in a sufficiently small neighborhood $\mathfrak{U}(Q)$ of Q, while, by Lemma 3.3, $\dim \Gamma(E+P) = d+1$. Hence $r_P{}^*\psi_0, \cdots, r_P{}^*\psi_d$ form a base of $\Gamma(E+P)$ for each $P \in \mathfrak{U}(Q)$. An arbitrary element $\phi \in \pi^{-1}(\mathfrak{U}(Q))$ can therefore be written in the form

$$\phi = \sum_{\nu=0}^{d} \zeta_\nu \cdot r_P{}^*\psi_\nu, \qquad \zeta_\nu \in \boldsymbol{C}. \tag{3.13}$$

In case $\phi \in \pi^{-1}(\mathfrak{U}(Q) \cap \mathfrak{U}(Q'))$, we have two representations

$$\phi = \sum_{\nu=0}^{d} \zeta_\nu \cdot r_P{}^*\psi_\nu = \sum_{\nu=0}^{d} \zeta_\nu{}' \cdot r_P{}^*\psi_\nu{}',$$

where $\psi_0', \cdots, \psi_d'$ are elements of $\Gamma(F)$ such that $r_{Q'}{}^*\psi_0', \cdots, r_{Q'}{}^*\psi_d'$ form a base of $\Gamma(E+Q')$. Obviously we have

$$r_P{}^*\psi_\nu = \sum_{\mu=0}^{d} \Theta_{\mu\nu}(P)\cdot r_P\psi_\mu' \tag{3.14}$$

and therefore

$$\zeta_\mu' = \sum_{\nu=0}^{d} \Theta_{\mu\nu}(P)\zeta_\nu, \tag{3.15}$$

where the coefficients $\Theta_{\mu\nu}(P)$ are holomorphic functions of $P \in \mathfrak{U}(Q)\cap\mathfrak{U}(Q')$ and $\det(\Theta_{\mu\nu}(P)) \neq 0$. Now we define the bundle structure on $\mathcal{B}$ by assigning, for each $Q \in \mathfrak{P}$, the product structure on $\pi^{-1}(\mathfrak{U}(Q))$ induced by the one-to-one correspondence

$$(P, \zeta_0, \zeta_1, \cdots, \zeta_d) \rightarrow \phi = \sum_{\nu=0}^{d} \zeta_\nu \cdot r_P{}^*\psi_\nu$$

between $\mathfrak{U}(Q)\times \boldsymbol{C}_{d+1}$ and $\pi^{-1}(\mathfrak{U}(Q))$. It follows from (3.13) and (3.15) that the bundle $\mathcal{B}$ thus defined is an analytic fibre bundle over $\mathfrak{P}$ whose fibre is the complex vector space $\boldsymbol{C}_{d+1}$ and whose structure group is the general linear group $GL(d+1, \boldsymbol{C})$.

Now replacing each fibre $\boldsymbol{C}_{d+1}$ of $\mathcal{B}$ by the corresponding projective space $\mathfrak{S}_d$, we derive from $\mathcal{B}$ a projective bundle Λ over $\mathfrak{P}$. More precisely, we first from the product space $\mathfrak{U}(Q)\times\mathfrak{S}_d$ for each point $Q\in\mathfrak{P}$, and then we derive the bundle Λ from the collection of the product spaces $\mathfrak{U}(Q)\times\mathfrak{S}_d$, $Q\in\mathfrak{P}$, by identifying $(P,\zeta)\in\mathfrak{U}(Q)\times\mathfrak{S}_d$ with $(P,\zeta')\in\mathfrak{U}(Q')\times\mathfrak{S}_d$ if and only if

$$\kappa\cdot\zeta_\mu' = \sum_{\nu=0}^{d}\Theta_{\mu\nu}(P)\zeta_\nu, \qquad \kappa\neq 0, \tag{3.16}$$

where $(\zeta_0, \cdots, \zeta_d)$ and $(\zeta_0', \cdots, \zeta_d')$ denote the homogeneous coordinates of ζ and ζ', respectively. The projective bundle Λ over $\mathfrak{P}$ is a non-singular projective variety,[25] since the base space $\mathfrak{P}$ is a non-singular projective variety. Let $\mathcal{B}^*$ be the subspace of $\mathcal{B}$ obtained by deleting the origin $(0, \cdots, 0)$ of each fibre $\boldsymbol{C}_{d+1}$ of $\mathcal{B}$. Then we have the canonical mapping

$$\phi = (P, \zeta_0, \cdots, \zeta_d) \rightarrow \lambda(\phi) = (P, \zeta)$$

of $\mathcal{B}^*$ onto Λ, where ζ is the point in $\mathfrak{S}_d$ with the homogeneous coordinate $(\zeta_0, \zeta_1, \cdots, \zeta_d)$. Since the divisor $D = (\phi)$ of $\phi = \sum\zeta_\nu r_P{}^*\psi_\nu$ is determined uniquely by $\lambda(\phi) = (P,\zeta)$, we set

$$D_\lambda = (\phi), \qquad \text{for } \lambda = \lambda(\phi). \tag{3.17}$$

[25] Kodaira [11], p. 42, Theorem 8.

In view of (3.1), the mapping $\phi \to D = (\phi)$ maps $\mathfrak{B}^*$ onto $\mathfrak{L} = \bigcup_{P \in \mathfrak{P}} |E + P|$, and moreover we have $(\phi') = (\phi)$ if and only if $\lambda(\phi') = \lambda(\phi)$. Consequently $\lambda \to D_\lambda$ gives a one-to-one correspondence between Λ and $\mathfrak{L}$. Moreover the correspondence $\lambda \to D_\lambda$ is *algebraic*, i.e., there exists an effective divisor $\mathfrak{D}$ on the product variety $V \times \Lambda$ such that

$$D_\lambda \times \lambda = (V \times \lambda) \cdot \mathfrak{D}, \qquad \text{for all } \lambda \in \Lambda. \tag{3.18}$$

This can be verified as follows: For any section $\psi: (z, P) \to \psi(z, P)$ belonging to $\Gamma(F)$, we denote by $\psi_{jQ}(z, P)$ the fibre coordinate of $\psi(z, P)$ over $U_j \times \mathfrak{U}(Q)$. Moreover we denote by $\{f_{jQkQ'}(z, P)\}$ the system of transition functions defining the bundle F with respect to the covering $\{U_j \times \mathfrak{U}(Q)\}$ of $V \times \mathfrak{P}$. We note that, for $P \in \mathfrak{U}(Q)$, the induced bundle $E + P = r_P{}^*F$ over V is defined by the system $\{f_{jQkQ}(z, P)\}$ with respect to the covering $\{U_j\}$ and that the mapping $\psi \to r_P{}^*\psi$ is simply given by

$$(r_P{}^*\psi)_j(z) = \psi_{jQ}(z, P). \tag{3.19}$$

Therefore it follows from (3.14) that

$$\psi_{\nu jQ}(z, P) = \sum_\mu \Theta_{\mu\nu}(P) \psi'_{\mu jQ}(z, P). \tag{3.20}$$

Now let $\mathfrak{O}_\mu$ be the open subset $\{\zeta \mid \zeta_\mu \neq 0\}$ of $\mathfrak{S}_d$. Then

$$R_{jQ\mu}(z, \lambda) = \sum_{\nu=0}^{d} (\zeta_\nu/\zeta_\mu) \psi_{\nu jQ}(z, P) \qquad \text{where } \lambda = (P, \zeta), \tag{3.21}$$

is a holomorphic function of (z, λ) defined on the open set

$$\mathcal{U}_{jQ\mu} = U_j \times \mathfrak{U}(Q) \times \mathfrak{O}_\mu \subset V \times \Lambda.$$

Moreover, since, by (3.20),

$$\psi_{\nu jQ}(z, P) = f_{jQkQ'}(z, P) \sum_\mu \Theta_{\mu\nu}(P) \psi'_{\mu kQ'}(z, P)$$

for $z \in U_j \cap U_k$, $P \in \mathfrak{A}(Q) \cap \mathfrak{A}(Q')$, we infer from (3.16) that

$$R_{jQ\mu}(z, \lambda)/R_{kQ'\nu}(z, \lambda) = f_{jQkQ'}(z, P) \sum_\tau \Theta_{\nu\tau}(P) \cdot (\zeta_\tau/\zeta_\mu), \qquad \text{in } \mathcal{U}_{jQ\mu} \cap \mathcal{U}_{kQ'\nu}. \tag{3.22}$$

The right hand side of (3.22) is obviously a non-vanishing holomorphic function on $\mathcal{U}_{jQ\mu} \cap \mathcal{U}_{kQ'\nu}$. Hence we can define an effective divisor $\mathfrak{D}$ on $V \times \Lambda$ by

$$\mathfrak{D} = (R_{jQ\mu}), \qquad \text{on each } \mathcal{U}_{jQ\mu}.$$

Clearly the fibre coordinate $\phi_j(z)$ of $\phi = \sum_{\nu} \zeta_\nu \cdot r_P{}^* \psi_\nu$ is written in the form

$$\phi_j(z) = \zeta_\mu \cdot R_{jQ\mu}(z, \lambda), \qquad \text{for } \lambda = (P, \zeta) \in \mathfrak{U}(Q) \times \mathfrak{D}_\mu.$$

Hence the divisor $D_\lambda = (\phi)$ is given by $D_\lambda = (R_{jQ\mu}(\ ,\lambda))$, on U_j. This proves (3.18). Thus we obtain the following

THEOREM 3.1. *Let D_0 be a divisor on V such that $|D_0|$ is not empty and let $\mathscr{B}$ be the set consisting of all effective divisors $D \sim D_0$ on V, where $\sim$ denotes the homology with coefficients in the integers. Assume that*

$$\text{(c)} \qquad c([D_0] - K) > 0, \qquad \textit{on } V,$$

where K is the canonical bundle over V. Then the set $\mathscr{B}$ forms an irreducible algebraic system of effective divisors parametrized by a non-singular projective vairety Λ in the sense that there exists a one-to-one analytic map $\lambda \to D_\lambda$ of Λ onto $\mathscr{B}$. The parameter variety Λ of $\mathscr{B}$ is an analytic fibre bundle over the Picard variety $\mathfrak{P}$ attached to V whose fibre is a projective space $\mathfrak{S}_d$ of dimension $d = \dim |D_0|$ and whose structure group is the projective transformation group acting on $\mathfrak{S}_d$. Moreover the canonical projection p of Λ onto $\mathfrak{P}$ is given by

$$p: \lambda \to p(\lambda) = [D_\lambda - D_0].$$

We note that the condition (c) in the above theorem is equivalent to

(c′) $\quad |m(D_0 - K)|$ is ample for sufficiently large integer m.

The complex line bundle $[\mathfrak{D}]$ over $V \times \Lambda$ associated with the divisor $\mathfrak{D}$ can be interpreted in the following manner: The subspace $\mathscr{B}^*$ of $\mathscr{B}$ is obviously a principal bundle over Λ whose structure group is $\boldsymbol{C}^*$. By replacing each fibre $\boldsymbol{C}^*$ of $\mathscr{B}^*$ by $\boldsymbol{C}$, we obtain from $\mathscr{B}^*$ a complex line bundle B over Λ. B is defined with respect to the covering $\{\mathfrak{U}(Q) \times \mathfrak{D}_\mu\}$ of Λ by the system $\{b_{Q\mu Q'\nu}(\lambda)\}$ of the functions

$$b_{Q\mu Q'\nu}(\lambda) = \zeta_\mu / \sum_{\tau} \Theta_{\nu\tau}(P) \zeta_\tau, \qquad \lambda = (P, \zeta).$$

Hence it follows from (3.22) that

$$\text{(3.23)} \qquad [\mathfrak{D}] = F - B,$$

where F in the right hand side denotes the complex line bundle over $V \times \Lambda$ induced by the bundle F over $V \times \mathfrak{P}$ in a canonical manner.

By a *continuous system* of divisors on V we shall mean an *irreducible* algebraic system of effective divisors on V. We say that two effective divisors

D and D' on V are *algebraically equivalent in the restricted sense*[26] and write $D \,|||\, D'$ if D and D' belong to the same continuous system. Moreover, for two (not necessarily effective) divisors D, D', we say that D and D' are *algebraically equivalent* if there exists a divisor X on V such that $D+X \,|||\, D'+X$. It is not difficult to show that *$D \,|||\, D'$ implies $D \sim D'$.* Conversely, *if $D \sim D'$, then the strict algebraic equivalence $D+X \,|||\, D'+X$ holds for any divisor X with sufficiently positive characteristic class $c(X)$*, as one readily infers from Theorem 3.1. Thus we obtain a well known theorem[27] to the effect that *the algebraic equivalence for divisors on a non-singular projective variety coincides with the homology equivalence.* A continuous system is said to be *complete* if it is not contained in a larger continuous system. A given effective divisor does not always determine uniquely the complete continuous system containing it. The continuous system $\mathscr{G}$ introduced in Theorem 3.1 is obviously complete. Moreover, $\mathscr{G}$ is the only complete continuous system containing the given divisor D_0.

4. Characteristic linear systems of complete continuous systems. Let C be an effective divisor on V and let $\mathfrak{M}$ be the set of all effective divisors $X \sim C$ on V. A "canonical parametrization" of this set $\mathfrak{M}$ can be introduced in the following manner[28]: Take a general hypersurface section L of V (in one of its ambient projective spaces) such that

$$c([C+L]-K) > 0. \tag{4.1}$$

Then, by Theorem 3.1, the set $\mathscr{G}$ of all effective divisors $D \sim C+L$ on V forms a complete continuous system parametrized by a projective bundle Λ over the Picard variety $\mathfrak{P}$ attached to V:

$$\mathscr{G} = \{D_\lambda \mid \lambda \in \Lambda\},$$

where the canonical projection p of Λ onto $\mathfrak{P}$ is given by

$$p:\ \lambda \to p(\lambda) = [D_\lambda - D_0], \qquad D_0 = C+L.$$

Now, letting $M = \{\lambda \mid D_\lambda > L\}$, we set $C_\lambda = D_\lambda - L$ for $\lambda \in M$. Then we infer readily that

$$\mathfrak{M} = \{C_\lambda \mid \lambda \in M\}, \tag{4.2}$$

while M is a (possibly reducible) algebraic subvariety of Λ. Thus the one-

[26] Zariski [19], p. 79.

[27] Lefschetz [16], Chap. IV; see also Igusa [7], p. 16.

[28] Weil [18], p. 887.

to-one mapping $\lambda \to C_\lambda$ of M onto $\mathfrak{M}$ gives a parametrization of $\mathfrak{M}$, which we call the *canonical parametrization* of $\mathfrak{M}$. In what follows we denote a point on M by μ instead of λ and write $\mu = \mu(X)$ if $X = C_\mu$.

LEMMA 4.1. *If C_o, $o \in M$, satisfies the condition*

$$H^1(V, \Omega(C_o)) = 0, \tag{a}$$

then $o = \mu(C_o)$ is a simple point of M. Moreover, for a sufficiently small neighborhood $\mathfrak{U}$ of $p(o)$ on $\mathfrak{P}$, $M \cap p^{-1}(\mathfrak{U})$ is an analytic sub-bundle of $\Lambda \cap p^{-1}(\mathfrak{U})$ in the sense that, for $P \in \mathfrak{U}$, $M \cap p^{-1}(P)$ is a linear sub-variety of the fibre $\mathfrak{S}_d = p^{-1}(P)$ depending holomorphically on P. The dimension h of the fibre $M \cap p^{-1}(P)$ is equal to $\dim | C_o |$.

Proof. Letting $B = [C_o]$ and $E = [D_o]$, $D_o = C_o + L$, we observe the exact sequence

$$0 \to \Omega(B+P) \to \Omega(E+P) \xrightarrow{r_L} \Omega((E+P)_L) \to 0,$$

where $P \in \mathfrak{P}$ and r_L denotes the restriction map to L. From the corresponding exact cohomology sequence

$$0 \to \Gamma(B+P) \to \Gamma(E+P) \to \Gamma((E+P)_L) \to H^1(V, \Omega(B+P)) \to \cdots$$

it follows that

$$0 \leqq \dim \Gamma(B+P) + \dim \Gamma((E+P)_L) - d - 1 \leqq \dim H^1(V, \Omega(B+P)),$$

where $d + 1 = \dim \Gamma(E+P)$. By hypothesis $\dim H^1(V, \Omega(B)) = 0$, while d is independent of P. Hence we get

$$\dim \Gamma(B+P) + \dim \Gamma((E+P)_L) \geqq \dim \Gamma(B) + \dim \Gamma(E_L).$$

On the other hand, $\dim \Gamma(B+P)$ and $\dim \Gamma((E+P)_L)$ are both upper semi-continuous functions in P, as Lemma 3.2 shows. Consequently, letting $h = \dim | C_o |$, we infer that there exists a neighborhood $\mathfrak{U}$ of $0 = p(o)$ on $\mathfrak{P}$ such that

$$\dim | C_o + P | = h, \qquad \text{for } P \in \mathfrak{U}. \tag{4.3}$$

It follows from (3.13) and (3.17) that, for any point

$$\lambda = (P, \zeta) \in p^{-1}(\mathfrak{U}) = \mathfrak{U} \times \mathfrak{S}_d,$$

the divisor D_λ is given by $D_\lambda = (\phi)$, $\phi = \sum_{\nu=0}^{d} \zeta_\nu \psi_\nu(P)$, where $\psi_\nu(P) = r^*{}_P \psi_\nu$. For each point $P \in \mathfrak{U}$, we have therefore

$$M \cap p^{-1}(P) = \{\mu \mid \mu = (P, \zeta), \quad \sum \zeta_\nu \psi_\nu(P)_L = 0\},$$

where $\psi_\nu(P)_L$ denotes the restriction of $\psi_\nu(P)$ to L. Since

$$P = p(\mu) = [D_\mu - D_o] = [C_\mu - C_o], \qquad \text{for } \mu \in M,$$

we have $|C_o + P| = \{C_\mu \mid \mu \in M \cap p^{-1}(P)\}$. Combined with (4.3) this shows that $\dim [M \cap p^{-1}(P)] = h$, for $P \in \mathfrak{U}$. Therefore we may assume that $\psi_{h+1}(P)_L, \cdots, \psi_d(P)_L$ are linearly independent and that

$$(4.4) \qquad \psi_\nu(P)_L = \sum_{\tau=h+1}^{d} A_{\nu\tau}(P)\psi_\tau(P)_L, \qquad (\nu = 0, 1, \cdots, h),$$

for $P \in \mathfrak{U}$, provided that $\mathfrak{U}$ is sufficiently small. It follows that the linear equation $\sum_{\nu=0}^{d} \zeta_\nu \psi_\nu(P)_L = 0$ is equivalent to the simultaneous equations

$$\zeta_\tau + \sum_{\nu=0}^{h} A_{\nu\tau}(P)\zeta_\nu = 0, \qquad (\tau = h+1, \cdots, d).$$

Consequently $M \cap p^{-1}(\mathfrak{U})$ is the subset of $p^{-1}(\mathfrak{U}) = \mathfrak{U} \times \mathfrak{S}_o$ consisting of all points of the form

$$(4.5) \qquad \mu = (P, \zeta), \qquad \zeta_\tau = -\sum_{\nu=0}^{h} A_{\nu\tau}(P)\zeta_\nu, \qquad P \in \mathfrak{U},$$

while it is easy to see that the coefficients $A_{\nu\tau}(P)$ determined by (4.4) are holomorphic functions of P. Thus $M \cap p^{-1}(\mathfrak{U})$ is an analytic sub-bundle of $p^{-1}(\mathfrak{U})$, q. e. d.

The canonical parametrization of $\mathfrak{M}$ *is unique*, i. e., it is independent of the choice of the auxiliary hypersurface section L of V. In fact, take any general hypersurface section L^* (in an arbitrary ambient projective space of V) satisfying (4.1) and construct the corresponding parametrization

$$\mathfrak{M} = \{C_{\mu^*} \mid \mu^* \in M^*\}, \qquad M^* \subset \Lambda^*.$$

Then the one-to-one mapping

$$\mu(X) \to \mu^*(X), \qquad X \in \mathfrak{M}$$

of M onto M^* is bi-regular in the sense that there exists a non-singular algebraic variety Σ and bi-regular mappings

$$\Phi: \Lambda \to \Sigma, \qquad \Phi^*: \Lambda^* \to \Sigma$$

of Λ, Λ^* *into* Σ such that $\Phi(\mu(X)) = \Phi^*(\mu^*(X))$, $X \in \mathfrak{M}$. This fact can be verified as follows: Consider the complete continuous system

$$\mathscr{E}^{**} = \{D^{**} \mid D^{**} \sim C + L + L^*, D^{**} > 0\}.$$

Since, by hypothesis, $c([C+L+L^*]-K) > c([C+L]-K) > 0$, the system $\mathscr{L}^{**}$ is parametrized by a non-singular algebraic variety Σ:

$$\mathscr{L}^{**} = \{D_\sigma^{**} \mid \sigma \in \Sigma\}.$$

Letting $\sigma(D^{**}) = \sigma$ for $D^{**} = D_\sigma^{**}$, we define a one-to-one mapping Φ of Λ into Σ by

$$\Phi : \lambda \to \Phi(\lambda) = \sigma(D_\lambda + L^*).$$

Clearly we have $\Phi(\Lambda) = \{\sigma \mid D_\sigma^{**} > L^*\}$. Moreover, for each $\sigma = \Phi(\lambda)$, $D_\lambda = D_\sigma^{**} - L^*$ satisfies $c([D_\lambda] - K) = c([C+L] - \mathrm{K}) > 0$, and therefore, by Theorem 1.2, $H^1(V, \Omega(D_\lambda)) = 0$. Hence we infer from the above Lemma 4.1 that $\Phi(\Lambda)$ is a non-singular subvariety of Σ. On the other hand, the one-to-one mapping Φ of Λ onto $\Phi(\Lambda) \subset \Sigma$ is analytic.[29] Consequently Φ is a bi-regular mapping[30] of Λ into Σ. Similarly the mapping

$$\Phi^* : \lambda^* \to \Phi^*(\lambda^*) = \sigma(D_{\lambda^*}^* + L)$$

is a bi-regular mapping of Λ^* into Σ. Moreover we have

$$\Phi(\mu(X)) = \sigma(X + L^* + L) = \Phi^*(\mu^*(X)), \qquad \text{for } X \in \mathfrak{M}.$$

We decompose the parameter variety M into the sum

$$M = M' \cup M'' \cup \cdots \cup M^{(l)} \cup \cdots$$

of irreducible components $M', M'', \cdots$ and set $\mathfrak{M}^{(l)} = \{C_\mu \mid \mu \in M^{(l)}\}$. Obviously we have the corresponding decomposition

$$\mathfrak{M} = \mathfrak{M}' \cup \mathfrak{M}'' \cup \cdots \cup \mathfrak{M}^{(l)} \cup \cdots.$$

Each component $\mathfrak{M}^{(l)}$ is a *complete continuous system*, i.e., an *irreducible algebraic system parametrized by* $M^{(l)}$. By the dimension of the complete continuous system $\mathfrak{M}^{(l)}$ we mean the dimension of the parameter variety $M^{(l)}$.

Suppose that $C = C_o$ is a non-singular prime divisor belonging to $\mathfrak{M}'$ and that $o = \mu(C)$ is a *simple point* of M'. Then the one-to-one mapping $\mu \to C_\mu$ is analytic in a neighborhood $\mathfrak{V}_o'$ of o on M' in the sense that there exists a divisor $\mathfrak{D}'$ on the open manifold $V \times \mathfrak{V}_o'$ such that

$$\mathfrak{D}' \cdot (V \times \mu) = C_\mu \times \mu, \qquad \mu \in \mathfrak{V}_o'. \tag{4.6}$$

In fact, letting $\{U_j\}$ be a sufficiently fine finite covering of V and letting $\mathfrak{U}_o$ be a neighborhood of o on Λ, we denote by $R_j(z, \lambda)$ the holomorphic function on $U_j \times \mathfrak{U}_o$ defining the divisor $\mathfrak{D}$ which determines the mapping $\lambda \to D_\lambda$

[29] Weil [18], p. 883.

[30] See footnotes 18 and 19.

of Λ onto $\mathcal{B}$. Moreover, let $w_j(z)$ be the holomorphic function on U_j defining the divisor L. Then, using the fact that L is a non-singular prime divisor, we infer readily that

$$S_j(z, \mu) = R_j(z, \mu)/w_j(z) \tag{4.7}$$

is a holomorphic function on $U_j \times \mathcal{V}_o'$ and that $b_{jk}(z, \mu) = S_j(z, \mu)/S_k(z, \mu)$ is a non-vanishing holomorphic function on $U_j \times \mathcal{V}_o' \cap U_k \times \mathcal{V}_o'$, provided that $\mathcal{V}_o' \subset \mathcal{U}_o$. Hence $\mathfrak{D}' = (S_j(z, \mu))$ is a well defined divisor on $V \times \mathcal{V}_o'$. Moreover (4.7) shows that for each $\mu \in \mathcal{V}_o'$, the function $S_j(z, \mu)$ in z has the divisor $D_\mu - L = C_\mu$. Consequently we obtain (4.6). We note that the parameter variety $\mathcal{V}_o'$ of the subset $\{C_\mu \mid \mu \in \mathcal{V}_o'\}$ of $\mathcal{M}'$ is *unique* in the following sense: Let $\mathcal{V}^*$ be an open complex manifold and let $\mu^* \to C_{\mu^*}$ be a one-to-one analytic map of $\mathcal{V}^*$ onto $\{C_\mu \mid \mu \in \mathcal{V}_o'\}$. Then $\mu^* \to \mu = \mu(C_{\mu^*})$ is a bi-regular map of $\mathcal{V}^*$ onto $\mathcal{V}_o'$. This is an immediate consequence of the fact that the one-to-one mapping $\mu^* \to \mu = \mu(C_{\mu^*})$ of $\mathcal{V}^*$ onto $\mathcal{V}_o'$ is analytic.[31]

Now the *characteristic linear system* of the complete continuous system $\mathcal{M}'$ on C may be defined as follows: We denote by $(\mu_1, \mu_2, \cdots, \mu_m)$ a system of local coordinates on M' with the center o and, for any "tangent vector" $u = (u_1, u_2, \cdots, u_m)$ of M' at o, we set

$$\partial_u = \sum_{\nu=1}^{m} u_\nu (\partial/\partial\mu_\nu).$$

Since the functions $S_j(z, o)$ vanish on $C = C_o$, we have

$$(\partial_u S_j(z, o))_C = b_{jk}(z, o)_C \cdot (\partial_u S_k(z, o))_C,$$

where $(\ )_C$ denotes restriction to C. On the other hand, the complex line bundle $B = [C] = [C_o]$ is defined by the system $\{b_{jk}(z, o)\}$. Hence

$$\phi_u \colon z \to \phi_u(z) = (z,\ \partial_u S_j(z, o)_C)$$

is a holomorphic section of the complex line bundle B_C over C, where $\partial_u S_j(z, o)_C$ stands for $(\partial_u S_j(z, o))_C$. Obviously we have

$$\phi_u = \sum u_\nu \phi_\nu, \quad \text{where } \phi_\nu(z) = (z, (\partial S_j(z, 0)/\partial\mu_\nu)_C)\,;$$

thus ϕ_u depends linearly on u. We associate with each tangent vector u such that $\phi_u \neq 0$ the effective divisor $\bar{C}_u = (\phi_u)$ on C. Then the set $\{\bar{C}_u\}$ consisting of all such divisors $\bar{C}_u$ forms a linear system contained in $|B_C|$. This

[31] See footnotes 18 and 19.

linear system $\{\bar{C}_u\}$ is called the *characteristic linear system*[32] *of* $\mathfrak{M}'$ *on* C. The characteristic linear system $\{\bar{C}_u\}$ is uniquely determined by $\mathfrak{M}'$ and C. In fact, it is obvious that $\{\bar{C}_u\}$ is uniquely determined by the divisor $\mathfrak{D}'$ and is independent of the choice of the system of local coordinates $(\mu_1, \cdots, \mu_m)$ and of the system $\{S_j(z,\mu)\}$, while $\mathfrak{D}'$ is determined uniquely by the mapping $\mu \to C_\mu$ of $\mathcal{V}_o'$ into $\mathfrak{M}'$. Hence the uniqueness of $\{\bar{C}_u\}$ follows from the uniqueness of the parameter variety $\mathcal{V}_o'$ mentioned above.

THEOREM 4.1. *Let C be an effective divisor on V satisfying*

$$\text{(a)} \qquad H^1(V, \Omega(C)) = 0.$$

Then C belongs to only one complete continuous system $\mathfrak{M}' = \{C_\mu \mid \mu \in M'\}$ and $\mu(C)$ is a simple point of the canonical parameter variety M' of $\mathfrak{M}'$. The dimension of the complete continuous system $\mathfrak{M}'$ is given by

$$\dim \mathfrak{M}' = \dim |C| + q, \tag{4.8}$$

where q is the dimension of the Picard variety $\mathfrak{P}$ attached to V. The system $\mathfrak{M}'$ contains all effective divisors $X \sim C$ on V satisfying $H^1(V, \Omega(X)) = 0$. If, moreover, C is a non-singular prime divisor, then the characteristic linear system of $\mathfrak{M}'$ on C is complete.

Proof. Let $\mathcal{B}$, Λ, p, $\mathfrak{M}$, M have the same meaning as above and let $C = C_o$, $o = \mu(C)$. By Lemma 4.1, there exists a neighborhood $\mathfrak{U}$ of $p(o)$ on $\mathfrak{P}$ such that $M \cap p^{-1}(\mathfrak{U})$ is an analytic sub-bundle of $p^{-1}(\mathfrak{U})$ whose fibre is a linear variety of dimension $h = \dim |C|$. It follows that there exists only one irreducible component, say M' of M which meets with $p^{-1}(\mathfrak{U})$ and that

$$M' \cap p^{-1}(\mathfrak{U}) = M \cap p^{-1}(\mathfrak{U}). \tag{4.9}$$

Hence $C = C_o$ is contained in $\mathfrak{M}'$ but not in $\mathfrak{M}^{(l)}$, $l \geqq 2$, and $o = \mu(C)$ is a simple point of M'. Moreover $M' \cap p^{-1}(\mathfrak{U})$ is an open manifold of dimension $h + q$ and therefore we obtain (4.8). Clearly the projection $p(M^{(l)})$ of each component $M^{(l)}$ is an algebraic subvariety of $\mathfrak{P}$. Hence it follows from (4.9) that $p(M') = \mathfrak{P}$ and that $p(M^{(l)})$, $l \geqq 2$, are proper subvarieties of $\mathfrak{P}$. Now, given an effective divisor $X \sim C$ satisfying $H^1(V, \Omega(X)) = 0$, we write $X = C_\mu$, $\mu \in M$. The above argument shows that the component $M^{(l)} \ni \mu$ has the projection $p(M^{(l)}) = \mathfrak{P}$. Hence $M^{(l)}$ coincides with M' and therefore $X \in \mathfrak{M}'$.

[32] It is easy to see that our definition of characteristic linear systems coincides with the classical one. See Zariski [19], p. 78.

Now, assuming that C is a non-singular prime divisor satisfying (a), we prove that the dimension of the characterstic linear system $\{\bar{C}_u\}$ of $\mathfrak{M}'$ on C is given by

$$\dim\{\bar{C}_u\} = h + q - 1, \qquad h = \dim|C|. \tag{4.10}$$

Obviously we may assume that

$$o = \mu(C) = (0, 1, 0, \cdots, 0)$$

in $p^{-1}(\mathfrak{U}) = \mathfrak{U} \times \mathfrak{S}_d$, where $\mathfrak{U} = \mathfrak{U}(0)$, $0 = p(o)$. It follows from (4.5) that a point $\mu \in \mathfrak{V}_o' \subset M' \cap \mathfrak{U}_o$ is written in the form

$$\mu = (P; 1, \eta_1, \cdots, \eta_h, \eta_{h+1}, \cdots, \eta_d),$$

where

$$\eta_\tau = -A_{0\tau}(P) - \sum_{\nu=1}^{h} A_{\nu\tau}(P)\eta_\nu, \qquad (\tau = h+1, \cdots, d).$$

Letting $P = \rho(t_1\bar{\alpha}_1 + \cdots + t_q\bar{\alpha}_q)$, for $P \in \mathfrak{U}$, we use

$$(\mu_1, \cdots, \mu_q, \mu_{q+1}, \cdots, \mu_{q+h}) = (t_1, \cdots, t_q, \eta_1, \cdots, \eta_h)$$

as the system of local coordinates on the neighborhood $\mathfrak{V}_o'$ of o on M'. The fibre coordinate $\psi_{\nu j}(z, P)$ of the section $\psi_\nu(P) = r_P{}^*\psi_\nu$ is given by $\psi_{\nu j}(z, P) = \psi_{\nu j0}(z, P)$, as (3.19) shows. Letting $f_{jQkQ'}(z, P)$ have the same meaning as in Section 3, we have therefore

$$\psi_{\nu j}(z, P) = f_{j0k0}(z, P)\psi_{\nu k}(z, P). \tag{4.11}$$

We infer from (4.4) that, for each $\nu = 0, 1, \cdots, h$,

$$\chi_{\nu j}(z, P) = w_j(z)^{-1}\{\psi_{\nu j}(z, P) - \sum_{\tau=h+1}^{d} A_{\nu\tau}(P)\psi_{\tau j}(z, P)\}$$

is a holomorphic function on $U_j \times \mathfrak{U}$. Moreover it follows from (4.11) that

$$\chi_{\nu j}(z, P) = b_{jk}(z, P) \cdot \chi_{\nu k}(z, P), \tag{4.12}$$

where

$$b_{jk}(z, P) = f_{j0k0}(z, P) \cdot w_k(z)/w_j(z). \tag{4.13}$$

Thus, for each $P \in \mathfrak{U}$,

$$\chi_\nu(P) : z \to (z, \chi_{\nu j}(z)), \qquad \nu = 0, 1, \cdots, h$$

are holomorphic sections of $B + P$; obviously these sections $\chi_0, \chi_1, \cdots, \chi_h$ are linearly independent and therefore form a base of the linear space $\Gamma(B + P)$ of dimension $h + 1$. Now, using the explicit form (3.21) of the functions $R_j(z, \lambda)$, we obtain

$$(4.14) \qquad S_j(z,\mu) = R_j(z,\mu)/w_j(z) = \chi_{0j}(z,P) + \sum_{\nu=1}^{h} \mu_{q+\nu}\chi_{\nu j}(z,P).$$

Clearly we have $S_j(z,\mu) = b_{jk}(z,P)\cdot S_k(z,\mu)$. For an arbitrary tangent vector $u = (u_1,\cdots,u_q, u_{q+1},\cdots,u_{q+h})$ of M' at o, we get, by (4.14),

$$(4.15) \qquad \partial_u S_j(z,o) = \partial_u \chi_{0j}(z,0) + \sum_{\nu=1}^{h} u_{q+\nu}\chi_{\nu j}(z,0)$$

where

$$\partial_u \chi_{0j}(z,0) = \sum_{k=1}^{q} u_k[\partial\chi_{0j}(z,P)/\partial t_k]_{t=0}.$$

In order to prove (4.10) it is sufficient to show that the holomorphic section

$$\phi_u\colon z \to \phi_u(z) = (z,\ \partial_u S_j(z,o)_C)$$

of B_C vanishes only if $u=0$. Assume that $\partial_u S_j(z,o)_C = 0$. Then, since C is the divisor of $S_j(z,o)$ on U_j, each ratio

$$s_j(z) = \partial_u S_j(z,o)/S_j(z,c)$$

is a holomorphic function on U_j. On the other hand, it follows from (4.12) that

$$(4.16) \qquad \mathfrak{f}_u(z) = \sum_{\nu=1}^{h} u_{q+\nu}\chi_{\nu j}(z,0)/\chi_{0j}(z,0)$$

is a well defined meromorphic function on V. Dividing (4.15) by $S_j(z,o) = \chi_{0j}(z,0)$, we get therefore

$$(4.17) \qquad s_j(z) - \partial_u \log \chi_{0j}(z,0) = \mathfrak{f}_u(z), \qquad \text{in } U_j.$$

Applying (2.3) to the function $f(\hat{z},\bar{\alpha};g,\bar{\delta})$ defining the bundle Ξ, we infer, in the same manner as in the proof of (2.5), that

$$f_{j0k0}(z,P) = e_{jk}(z)\kappa_j\kappa_k^{-1}\cdot \exp\int_{z(j)}^{z(k)}\bar{\alpha},$$

where $\{e_{jk}(z)\}$ is the system of transition functions defining the bundle $E = B + [L]$ and where

$$\kappa_j = \exp\int_{\hat{o}}^{\hat{z}(j)} \varpi^*\bar{\alpha}.$$

Inserting this into (4.13), we obtain

$$b_{jk}(z,P) = b_{jk}(z,0)\,\kappa_j\kappa_k^{-1}\exp\int_{z(j)}^{z(k)}\bar{\alpha}.$$

Using (4.12) we get therefore

$$\partial_u \log \chi_{0j}(z,0) - \partial_u \log \chi_{0k}(z,0) = \iota_j - \iota_k + \int_{z(j)}^{z(k)} \bar{\alpha}_u,$$

where $\bar{\alpha}_u = \sum_{k=1}^{q} u_k \bar{\alpha}_k$ and where $\iota_j = \int_{\hat{o}}^{\hat{z}(j)} \varpi^* \bar{\alpha}_u$. Combining this with (4.17) we obtain

$$s_j(s) - \iota_j - \int_{z(j)}^{z} \bar{\alpha}_u = s_k(z) - \iota_k - \int_{z(k)}^{z} \bar{\alpha}_u.$$

Thus $\rho(z) = s_j(z) - \iota_j - \int_{z(j)}^{z} \bar{\alpha}_u$ is a well defined function of class C^∞ on V and $d\rho(z) = ds_j(z) - \bar{\alpha}_u$. This implies that $ds_j = ds_k$ is a well defined holomorphic 1-form on V. Consequently, we have $ds_j(z) = 0$ and $\bar{\alpha}_u = 0$. The equality $\bar{\alpha}_u = 0$ implies that $u_1 = u_2 = \cdots = u_q = 0$. Hence we get $\partial_u \log \chi_{0j}(z,0) = 0$ and therefore, by (4.17), $d\mathfrak{f}_u(z) = ds_j(z) = 0$. Thus $\mathfrak{f}_u(z) = c_0$ is a constant on V and consequently by (4.16),

$$c_0 \chi_{0j}(z,0) + \sum_{\nu=1}^{h} u_{q+\nu} \chi_{\nu j}(z,0) = 0.$$

This implies that $u_{q+1} = u_{q+2} = \cdots = u_{q+h} = 0$. Thus we see that ϕ_u vanishes only if $u = 0$. This completes the proof of (4.10).

The characteristic linear system $\{\bar{C}_u\}$ of $\mathfrak{M}'$ on C is a subsystem of the complete linear system $|B_C|$. In order to prove that $\{\bar{C}_u\}$ coincides with $|B_C|$, we observe the exact sequence

$$0 \to \Omega \to \Omega(B) \xrightarrow{r_C} \Omega(B_C) \to 0.$$

The corresponding exact cohomology sequence

$$0 \to \boldsymbol{C} \to \Gamma(B) \to \Gamma(B_C) \to H^1(V, \Omega) \to \cdots$$

shows that

$$\dim |B_C| \leqq \dim |B| - 1 + \dim H^1(V, \Omega),$$

while we have $H^1(V,\Omega) \cong \bar{\mathfrak{A}}$, and so $\dim H^1(V,\Omega) = \dim \mathfrak{P} = q$. Consequently we obtain the *inequality*

$$\dim |B_C| \leqq \dim |C| + q - 1. \tag{4.18}$$

Comparing this with (4.10), we infer immediately that $\{\bar{C}_u\} = |B_C|$. Thus the characteristic linear system $\{\bar{C}_u\}$ of $\mathfrak{M}'$ on C is complete.

We note that a divisor C on V satisfies the above condition (a) if

$$c([C] - K) > 0 \tag{c}$$

holds. This is an immediate consequence of Theorem 1.2. Moreover it follows from a remark in Section 1 that the condition (c) is equivalent to

(c′) $|m(C-K)|$ is ample for sufficiently large integer m.

Another remark is that the condition (a) is actually weaker than (c). This can be easily seen by observing the simplest case where V is an *algebraic curve* of genus p. In fact, a general effective divisor C of degree $d \geqq p$ on V satisfies (a), while (c) holds for C if and only if $d > 2p-2$.

THEOREM 4.2. *Let* $\mathfrak{M}' = \{C_\mu \mid \mu \in M'\}$ *be a complete continuous system canonically parametrized by* M'. *Assume that*

$$H^1(V, \Omega(C_\mu)) = 0, \qquad \textit{for all } C_\mu \in \mathfrak{M}'.$$

Then the parameter variety M' *is an analytic fibre bundle over the Picard variety* $\mathfrak{P}$ *of* V *whose fibre is the projective space* $\mathfrak{S}_h$ *of dimension* $h = \dim |C_\mu|$ *and whose structure group is the projective transformation group. The canonical projection of* M' *onto* $\mathfrak{P}$ *is given by*

$$p: \mu \to p(\mu) = [C_\mu - C_o], \qquad o \in M'.$$

Moreover $\mathfrak{M}'$ *consists of all effective divisors* $X \sim C_o$, $C_o \in \mathfrak{M}'$, *on* V.

Proof. Let $\mathcal{B}$, Λ, p, $\mathfrak{M}$, M have the same meaning as above. We infer from Lemma 4.1 that, for each point $P = p(\mu)$, $\mu \in M'$, there exists a neighborhood $\mathfrak{U}(P)$ of P on $\mathfrak{P}$ such that $M \cap p^{-1}(\mathfrak{U}(P))$ is an analytic subbundle of $p^{-1}(\mathfrak{U}(P))$ whose fibre is a linear variety of dimension $h = \dim |C_\mu|$. It follows that

$$M' \cap p^{-1}(\mathfrak{U}(P)) = M \cap p^{-1}(\mathfrak{U}(P))$$

and therefore $p(M') \supset \mathfrak{U}(P)$, $P \in p(M')$, while $p(M'$ is a compact subset of $\mathfrak{P}$. Hence $p(M')$ coincides with $\mathfrak{P}$ and consequently M' is an analytic fibre bundle over $\mathfrak{P}$ whose fibre is $\mathfrak{S}_h$ and whose structure group is the projective transformation group. Moreover M' coincides with M and therefore $\mathfrak{M}'$ coincides with $\mathfrak{M}$, q. e. d.

In the special case where V is an *algebraic curve* of genus p, the *index of speciality* $i\,|C_\mu| = \dim H^1(V, \Omega(C_\mu))$ vanishes if the degree of the divisor C_μ is greater than $2p-2$. Therefore our result reduces in this case to a theorem of Chow[33] to the effect that the continuous system of all effective divisors of degree $d > 2p-2$ on the curve V is a projective bundle over the Jacobian variety attached to V. Thus our Theorem 4.2 may be regarded as a generalization of the above theorem of Chow to higher dimensional varieties.

PRINCETON UNIVERSITY AND THE INSTITUTE FOR ADVANCED STUDY.

[33] Chow [5].

REFERENCES.

[1] W. L. Baily, "On the quotient of an analytic manifold by a group of analytic homeomorphisms," *Proceedings of the National Academy of Sciences*, vol. 40 (1954), pp. 804-808.

[2] ———, *On the quotient of a complex analytic manifold by a discontinuous group of complex analytic self-homeomorphisms*. Thesis, Princeton (1955).

[3] S. Bochner and W. T. Martin, *Several complex variables*, Princeton, 1948.

[4] W. L. Chow, "On Picard varieties," *American Journal of Mathematics*, vol. 74 (1952), pp. 895-909.

[5] ———, "The Jacobian variety of an algebraic curve, *ibid.*, vol. 76 (1954), pp. 453-476.

[6] ———, "On complex analytic varieties, *ibid.*, to appear.

[7] J. Igusa, "On Picard varieties attached to algebraic varieties," *ibid.*, vol. 74 (1952), pp. 1-22.

[7a] ———, "On some problems in abstract algebraic geometry," to appear in the *Proceedings of the National Academy of Sciences*.

[8] K. Kodaira, "Some results in the transcendental theory of algebraic varieties," *Annals of Mathematics*, vol. 59 (1954), pp. 86-134.

[9] ———, "On cohomology groups of compact analytic varieties with coefficients in some analytic faisceaux," *Proceedings of the National Academy of Sciences*, vol. 39 (1953), pp. 865-868.

[10] ———, "On differential-geometric method in the theory of analytic stacks," *ibid.*, vol. 39 (1953), pp. 1268-1273.

[11] ———, "On Kähler varieties of restricted type," *Annals of Mathematics*, vol. 60 (1954), pp. 28-48.

[12] ——— and D. C. Spencer, "Groups of complex line bundles over compact Kähler varieties," *Proceedings of the National Academy of Sciences*, vol. 39 (1953), pp. 868-872.

[13] ——— and D. C. Spencer, "Divisor class groups on algebraic varieties," *ibid.*, vol. 39 (1953), pp. 873-877.

[14] ——— and D. C. Spencer, "On the varation of almost-complex structure," Mimeographed note, Princeton (1954).

[15] S. Lefschetz, "On certain numerical invariants of algebraic varieties with application to abelian varieties," *Transactions of the American Mathematical Society*, vol. 22 (1921), pp. 327-482.

[16] ———, *L'Analysis situs et la géometrie algébrique*, Paris, 1924.

[17] F. Severi, "Sul teorema fondamentale dei sistemi continui di curve sopra una superficie algebrica," *Annali di Matematicae*, vol. 23 (1944), pp. 149-181.

[18] A. Weil, "On Picard varieties," *American Journal of Mathematics*, vol. 74 (1952), pp. 865-894.

[19] O. Zariski, *Algebraic Surfaces*, Ergebnisse der Mathematik und ihrer Grenzgebiete, vol. 3, No. 5, 1935.

[20] G. Zappa, "Sull'esistenza, sopra le superficie algebriche, di sistemi continui completi infiniti, la cui curva generica e a serie caratteristica incompleta," *Pont. Acad. Sci. Acta*, vol. 9 (1945), pp. 91-93.

of F_0+F_1, we have only to add a suitable hypersurface to F_0+F_1 so that we get a hypersurface F of order m.

Now, coming back to our original problem, we shall prove the following assertion:

LEMMA 1. *Let G be a commutative group variety which is defined, together with its group law, over a field K; let t be a point of G. Then, $x\cdot y=x+y+t$ is a group law in G over K if and only if t is rational over K. Moreover, if G is an Abelian variety, every group law in G over K can be so obtained.*

Proof. The first assertion is trivial. Assume that G is an Abelian variety over K. Let ϕ be a new group law in G over K. Then, we can find an automorphism λ of G over K with respect to the old group law in G, and a rational point t of G over K such that $\phi(x,y)=\lambda(x+y)+t$ for x and y on G [6, p. 34]. The associativity of the law implies $\lambda^2=\lambda$, whence $\lambda=\delta_G$, the identity automorphism of G. q. e. d.

We pick an arbitrary rational point t of J over $k(u)$; let t' be the specialization of t over the specialization $u\rightarrow a$ with reference to k. Then, by Theorem 1 and by Lemma 1 we can introduce group laws in J and in $(J')_0$ over $k(u)$ and over k, respectively, such that $-t$ and $-t'$ become new neutral elements. Henceforth, we shall denote $-t$ by 0 and $-t'$ by $0'$. We note that the group laws are compatible with specialization [2].

Let n be a positive integer which is not divisible by the characteristic p of the universal domain. Then, δ being the identity automorphism of J, we denote by T the graph of $n\cdot\delta$ in the product $J\times J$. As we know, if x is an arbitrary point of J, the cycle inverse $(n\cdot\delta)^{-1}(x)=\mathrm{pr}_1[T\cdot(J\times x)]$ is defined, and consists of n^{2g} distinct points. In fact, this is a general theorem which holds for any Abelian variety [5, p. 127]. Similarly, let δ' be the identity automorphism of $(J')_0$, and let $(T')_0$ be the graph of $n\cdot\delta'$ in $(J')_0\times(J')_0$. With these notations, we shall prove the following lemma:

LEMMA 2. *As an abstract group $(J')_0$ splits over G_m,* i. e., *$(J')_0=G_m\times((J')_0/G_m)$. Moreover, the cycle inverse*

$$(n\cdot\delta')^{-1}(x)=\mathrm{pr}_1[(T')_0\cdot((J')_0\times x)]$$

is defined for every point x of $(J')_0$, and consists of n^{2g-1} distinct points.

Proof. The first assertion is a consequence of a general theorem on Abelian groups, because G_m is "infinitely divisible."[2] Then, the facts that

[2] A precise statement of the theorem is the following: *Let G be an Abelian group, and let H be it subgroup which is infinitely divisible. Then, the extension G of H splits over H.*

the cycle inverse $(n\cdot\delta')^{-1}(x)$ is defined for every point x of $(J')_0$ and that the number of distinct points in it is equal to n^{2g-1} follow from the corresponding properties of the factors of $(J')_0$. It is, therefore, sufficient to show that the homomorphism $n\cdot\delta'$ is separable. A usual proof of this for any commutative group variety G, and for its n-times the identity automorphism δ_G is the following: We know that the ring of endomorphisms of G has a representation on th space of forms of Maurer-Cartan. The endomorphism $n\cdot\delta_G$ is represented by the n-times the unit matrix, and none of the forms of Maurer-Cartan becomes zero by multiplication of n. Since these forms generate the space of linear differential forms over the function-field, if x is a generic point of G over one of its fields of definition, say K, there is no derivation of $K(x)$ over $K(n\cdot x)$ except, of course, the trivial one. Hence $K(x)$ is separably algebraic over $K(n\cdot x)$. q. e. d.

Now, let T' be the closure of $(T')_0$ with respect to the Zariski topology. Then, from Lemma 2 we get $\mathrm{pr}_2(T')=n^{2g-1}\cdot J'$. On the other hand, by our previous remark we have $\mathrm{pr}_2(T)=n^{2g}\cdot J$. Therefore, if we specialize T over the specialization $u\to a$ with reference to k, we get T' and something more. The next lemma concerns about this:

LEMMA 3. *The specialization, say $\bar{T}$, of T over the specialization $u\to a$ with reference to k is of the form $\bar{T}=T'+T''$. Here, T'' is a positive cycle of the product $P^N\times P^N$ each component of which has a singular subvariety of J' as its projection on the first factor.*

Proof. Let $x'\times y'$ be a generic point over k of some component, say W, of $\bar{T}$. Then, we can find a generic point $x\times n\cdot x$ of T over $k(u)$ such that $x'\times y'$ is a specialization of $x\times n\cdot x$ over the specialization $u\to a$ with reference to k. If x' is a simple point of J', since the group laws are compatible with specialization, we have $y'=n\cdot x'$. In other words, if the projection of W on the first factor is not contained in the singular locus of J', then, W coincides with T'. Since $\mathrm{pr}_1(T')=J'$ is the specialization of $\mathrm{pr}_1(T)=J$ over the specialization $u\to a$ with reference to k, we see that $\bar{T}$ does contain T', and in fact only once. q. e. d.

It is a simple matter to determine all components of T''. We get just what is expected. However, since we shall not make use of such information, we do not discuss it here.

Now, let r be a rational point of J over $k(u)$, and let r' be the specialization of r over the specialization $u\to a$ with reference to k. Then, by Theorem 1 the point r' is simple on J'. We shall denote the 0-cycles $(n\cdot\delta)^{-1}(r)$ and $(n\cdot\delta')^{-1}(r')$ by w and w'. We remark that w is rational over $k(u)$, hence it has a unique specialization over the specialization $u\to a$ with reference to k. With this in mind, we shall prove the following theorem:

THEOREM 2. *The specialization, say* $\bar{w}$, *of* w *over the specialization* $u \to a$ *with reference to* k *is of the form* $\bar{w} = w' + w''$. *Here,* w'' *is a positive cycle of* P^N *each component of which is a multiple point of* J'.

Proof. We first note that w and w' are defined, respectively, as $\mathrm{pr}_1[T \cdot (J \times r)]$ and $\mathrm{pr}_1[(T')_0 \cdot ((J')_0 \times r')]$. Let s' be an arbitrary simple point of J'. Then, $s' \times r'$ is not contained in any component of T'' of Lemma 3. Since $(T' + T'', J', r')$ is the specialization of (T, J, r) over the specialization $u \to a$ with reference to k, by the "principle of conservation of number"[3] $s' \times r'$ appears in $\bar{w} \times r'$ and in $T' \cdot (J' \times r')$ exactly the same number of times. However, we have $T' \cdot (J' \times r') = (T')_0 \cdot ((J')_0 \times r')$, the intersection-products being taken on $J' \times J'$ and on $(J')_0 \times (J')_0$. This is nothing but $w' \times r'$. q. e. d.

Let $k[[u - a]]$ be the completion of the specialization-ring of a in $k(u)$, and let $k((u - a))$ be its quotient-field. We note that $k[[u - a]]$ is a ring of formal power-series in one letter over k. In the following, we consider $k(u)$ as a subfield of the "abstract field" $k((u - a))$. Let I be an isomorphism over $k(u)$ of the algebraic closure of $k(u)$ into some, but fixed, algebraic closure of $k((u - a))$. Let (w) be an ordered set of points in w, and let $I(w)$ be the transform of (w) by I. Then, the specialization of $I(w)$ at the center of $k[[u - a]]$ is an ordered set of points in $\bar{w}$, because it is a specialization of (w) over the specialiaztion $u \to a$ with reference to k. In particular, (w) contains n^{2g-1} points, say $(s_1, s_2, \cdots)$, such that the specialization of $I(s_1, s_2, \cdots)$ at the center of $k[[u - a]]$ is an ordered set of points of w'. Then, by generalized "Hensel's lemma" each $I(s_i)$ is rational over $k((u - a))$ (cf. [5], Theorem 2, p. 57). We note that the set $(s_1, s_2, \cdots)$ is uniquely determined up to a permutation of elements by I. Now, assume that r and r' are the neutral elements 0 and $0'$ of J and $(J')_0$. Then, the points $s_1, s_2, \cdots$ form an Abelian group, which is a direct product of $2g - 1$ cyclic groups of order n. We call it the *group of invariant points of order* n *along* I. In fact, each $I(s_i)$ is "invariant" by any automorphism of the algebraic closure of $k((u - a))$ over $k((u - a))$. On the other hand, this group contains a particular set of n points, say $t_1, t_2, \cdots$, such that $I(t_i)$ is specialized to a point of G_m at the center of $k[[u - a]]$. We note that the

[3] This famous principle has various formulations. Here, it is used with the following content: *Let* X *and* Y *be two positive cycles on a variety* V *in a projective space such that* $\dim(X) + \dim(Y) = \dim(V)$; *let* (V', X', Y') *be a specialization of* (V, X, Y) *over a field* k, *and assume that* V' *is a variety. If the intersection-products* $X \cdot Y$ *and* $X' \cdot Y'$ *are defined on* V *and on* V', *a simple point* P' *of* V' *appears* $i(X' \cdot Y', P'; V')$-*times in any specialization of* $X \cdot Y$ *over the specialization* $(V, X, Y) \to (V', X', Y')$ *with reference to* k.

points $t_1, t_2, \cdots$ form a cyclic group of order n, and we call it the *group of vanishing points of order n along I*. In fact, G_m is the set of those points of $(J')_0$ which "vanish" by projection to the Jacobian variety J^*!

3. Non-invariant points. In the previous section, we introduced invariant and vanishing points. We shall now examine "non-invariant" points; the following is our main theorem:

THEOREM 3. *Let r be a rational point of J over $k(u)$; let x and y be two points of J such that $n \cdot x = n \cdot y = r$, and such that $I(x)$ and $I(y)$ are conjugate over $k((u-a))$. Then, $x-y$ is a vanishing point of order n along I.*

Proof. Since the theorem is trivial for $x=y$, we shall exclude this case from the beginning. Then, by Theorem 2 the specialization, say x', of $I(x)$ at the center of $k[[u-a]]$ is a multiple point of J'. Also, by Hensel's lemma (x', x') is the specialization of $(I(x), I(y))$ at the center of $k[[u-a]]$. Now, take an integer m such that the trace $\mathfrak{L}_m$ on C of the complete linear system of hypersurfaces of order m in the ambient space is complete, and of dimension equal to $r = \mathrm{ord}(C) \cdot m - g$. We also assume that the trace $\mathfrak{L}_m'$ on C' of the same linear system of hypersurfaces is of dimension r [8]. If m satisfies these conditions, we call it to be sufficiently large. We pick rational divisors $\mathfrak{a}$ and $\mathfrak{s}$ of C of degree d over $k(u)$ which correspond to 0 and r as remarked next to Theorem 1; put $\mathfrak{a} = \mathfrak{a}_0 - \mathfrak{a}_1$ and $\mathfrak{s} = \mathfrak{s}_0 - \mathfrak{s}_1$. We then take a hypersurface of a sufficiently large order, say m, for $\mathfrak{s}_0 + (n-1) \cdot \mathfrak{a}_0$ as stated in the Supplement of Theorem 1. If $\mathfrak{r}_1$ is the sum of $\mathfrak{s}_1 + (n-1) \cdot \mathfrak{a}_1$ and the residual intersection on C, then, $\mathfrak{r}_1$ is rational over $k(u)$, and the specialization, say $\mathfrak{r}_1'$, of $\mathfrak{r}_1$ over the specialization $u \rightarrow a$ with reference to k is free from Q. Consider the residual linear system $\mathfrak{L}_m - \mathfrak{r}_1$. The set of positive divisors of degree d on C such that their n-times belong to $\mathfrak{L}_m - \mathfrak{r}_1$ splits into n^{2g} distinct complete linear systems on C, and these linear systems correspond in a one-to-one way to points of J such as x and y. Let $|\mathfrak{m}|$ be the complete linear system which corresponds to x. We may assume that $\mathfrak{m}$ is algebraic over $k(u)$, and such that the specializations of $\mathfrak{m}$ over the specialization $(u, x) \rightarrow (a, x')$ with reference to k contain Q once and only once. In fact, let $\mathfrak{M}$ and $\mathfrak{M}'$ be the varieties which have x and x' as their Chow points. Then, we can find a subspace of co-dimension $d-g$, in the ambient space of $\mathfrak{M}$ and $\mathfrak{M}'$, which intersects properly with $\mathfrak{M}$ and $\mathfrak{M}'$. In addition, we can assume that the subspace is defined over k, and also that it does not intersect with the closed subset of $\mathfrak{M}'$ whose points represent those cycles which contain Q at least twice [2]. If $\mathfrak{m}$ is one of the divisors of C

which correspond to the intersection of $\mathfrak{M}$ and the subspace, $\mathfrak{m}$ satisfies our requirements. Now, there are many hypersurfaces of order m which intersect C at $n\cdot\mathfrak{m}+\mathfrak{r}_1$. However, since C and C' have the same arithmetic genus, we can pick one, say F_1, which is algebraic over $k(u)$, and such that F_1 has a unique specialization over any specialization $(u,\mathfrak{m})\rightarrow(a,\mathfrak{m}')$ with reference to k, and such that this specialization does not contain C'. In fact, the complete linear system of hypersurfaces of order m in the ambient space of C and C' gives rise a regular rational map of this projective space into another projective space, say P. Let P^r and P'^r be the smallest subspaces of P which contain the images of C and C', respectively. Take a subspace P^t of P of co-dimension $r+1$ which spans P with P^r, and also with P'^r. Since P is defined over the prime field, we can assume P^t to be algebraic over the prime field. On the other hand, since $n\cdot\mathfrak{m}+\mathfrak{r}_1$ is rational over $k(u,\mathfrak{m})$, the hyperplane P^{r-1} of P^r, which corresponds to $n\cdot\mathfrak{m}+\mathfrak{r}_1$, is purely inseparable over $k(u,\mathfrak{m})$. If F_1 is the hypersurface of the original projective space which is represented by the join of P^{r-1} and P^t, then, F_1 satisfies our demands. Let F' be the specialization of $I(F_1)$ at the center of $k[[u-a]]$. We note that $I(F_1)$ has a meaning, because F_1 is algebraic over $k(u)$. Let $(I(y),\bar{\mathfrak{m}},\bar{F}_1)$ be a conjugate of $(I(x),I(\mathfrak{m}),I(F_1))$ over $k((u-a))$. Then, we can find a positive divisor $\mathfrak{n}$ of degree d on C and a hypersurface F_2 of order m such that $\bar{\mathfrak{m}}=I(\mathfrak{n})$ and $\bar{F}_1=I(F_2)$. Since $(\mathfrak{n},F_2)$ is a conjugate of $(\mathfrak{m},F_1)$ over $k(u)$, we see that F_2 intersects C at $n\cdot\mathfrak{n}+\mathfrak{r}_1$. Since $(y,\mathfrak{n})$ is a conjugate of $(x,\mathfrak{m})$ over $k(u)$, we see that $\mathfrak{n}$ corresponds to y. Also, by Hensel's lemma $(\mathfrak{m}',\mathfrak{m}')$, say, and $(F',F')$are the specializations of $(I(\mathfrak{m}),I(\mathfrak{n}))$ and $(I(F_1),I(F_2))$ at the center of $k[[u-a]]$. This will become a key point in our later argument. Now, take a hypersurface, say G_1, of a sufficiently large order m', which passes through $\mathfrak{m}+\mathfrak{a}_0$. We take G_1 to be algebraic over $k(u)$, and such that the specialization, say G_1', of $I(G_1)$ at the center of $k[[u-a]]$ does not contain C'. Also, we assume that G_1' contains Q only twice. In fact, these can be managed, more or less, by similar considerations as before. Let $\mathfrak{r}_2$ be the sum of $\mathfrak{a}_1$ and the residual intersection of G_1 on C. Then, the residual linear system $\mathfrak{L}_{m'}-(\mathfrak{n}+\mathfrak{r}_2)$ exists, and it is the complete linear system which corresponds to $x-y$. We shall show that the specialization, say $(x-y)'$, of $I(x-y)$ at the center of $k[[u-a]]$ is a simple point of J'. Otherwise, we can pick $\mathfrak{p}$ from $\mathfrak{L}_{m'}-(\mathfrak{n}+\mathfrak{r}_2)$ such that $\mathfrak{p}$ is algebraic over $k(u)$, and such that the specialization $\mathfrak{p}'$ of $I(\mathfrak{p})$ at the center of $k[[u-a]]$ contains Q once and only once. Let G_2 be a hypersurface of order m' which is algebraic over $k(u)$, and which intersects C at $\mathfrak{p}+\mathfrak{n}+\mathfrak{r}_2$. We may assume that the specialization, say G_2', of $I(G_2)$ at the center of $k[[u-a]]$ does not contain

C'. Finally, let H_1 be a hypersurface of a sufficiently large order, say m'', for $n\cdot\mathfrak{a}_0$ as stated in the Supplement of Theorem 1. If $\mathfrak{r}_3$ is the sum of $n\cdot\mathfrak{a}_1$ and the residual intersection on C, then, $\mathfrak{r}_3$ is rational over $k(u)$, and the specialization, say $\mathfrak{r}_3'$, of $\mathfrak{r}_3$ over the specialization $u\to a$ with reference to k is free from Q. Let H_2 be a hypersurface of order m'' which is algebraic over $k(u)$, and which intersects C at $n\cdot\mathfrak{p}+\mathfrak{r}_3$. We may assume that the specialization, say H_2', of $I(H_2)$ at the center of $k[[u-a]]$ does not contain C'. Then, by our construction two hypersurfaces $F_1+n\cdot G_2+H_1$ and $F_2+n\cdot G_1+H_2$ have the same intersection with C. Therefore, the specializations $F'+n\cdot G_2'+H_1$ and $F'+n\cdot G_1'+H_2'$ have the same intersection with C'. More precisely, if C^* is the normalization of C' over k, and if f is the corresponding regular rational map of C^* into the ambient space of C', we have $f^{-1}(F'+n\cdot G_2'+H_1)=f^{-1}(F'+n\cdot G_1'+H_2')$, and hence $n\cdot f^{-1}(G_2')+f^{-1}(H_1)=n\cdot f^{-1}(G_1')+f^{-1}(H_2')$. Let Q_1^* and Q_2^* be the points of C^* which correspond to Q. Let c_1 and c_2 be the coefficients of Q_1^* and Q_2^* in the reduced expression of the divisor of C^* just obtained. Since H_1 does not pass through Q, if we use the first expression, we conclude that both c_1 and c_2 are multiples of n. On the other hand, H_2' intersects C' at $n\cdot\mathfrak{p}'+\mathfrak{r}_3'$, and $\mathfrak{p}'$ contains Q once and only once, while $\mathfrak{r}_3'$ is free from Q. Since $f^{-1}(H_2')$ must contain both Q_1^* and Q_2^*, we conclude from the second expression that neither c_1 nor c_2 is divisible by n. Thus, we get a contradiction, and this contradiction is derived from the assumption that $(x-y)'$ is a multiple point of J'. In other words, we have shown that $(x-y)'$ is a simple point of J'. Then, our previous argument goes through under the modification that $\mathfrak{p}'$ is free from Q. Since $\mathfrak{m}'$ contains Q only once and G_1' does only twice, we conclude that $f^{-1}(G_2')-f^{-1}(G_1')=\mathfrak{p}^*-\mathfrak{a}^*$. Here, $\mathfrak{p}^*$ and $\mathfrak{a}^*$ correspond to $\mathfrak{p}'$ and the specialization, say $\mathfrak{a}'$ of $\mathfrak{a}$ over the specialization $u\to a$ with reference to k. Since the left side is linearly equivalent to zero on C^*, so is also the right side, i.e., $(x-y)'$ belongs to G_m. Therefore, by definition $x-y$ is a vanishing point of order n along I. q.e.d.

Remark. If the linear equivalence class of $Q_1^*-Q_2^*$ on C^* is of order at least equal to n, we can prove the following assertion: *Let x and y be two points of J such that $n\cdot x=n\cdot y=r$, and such that (x',x'), say, is a specialization of (x,y) over the specialization $u\to a$ with reference to k. Then, any specialization of $x-y$ over this specialization is a point of G_m.* A proof of this assertion can be extracted from the proof we gave for Theorem 3.

4. A bilinear relation. In this section, we shall show that invariant points and vanishing points are connected by a bilinear relation. Let U be,

for a moment, an arbitrary complete curve. Then, for any numerical function f on U, we can define a divisor (f), and a divisor in the sense of valuation theory. The latter is defined as the regular image on U of the divisor of the function induced by f on the non-singular model of U. On the other hand, if $\sum_{\alpha} a_{\alpha} \cdot P_{\alpha}$ is a reduced expression of a divisor $\mathfrak{a}$ on U, and if (f) does not contain any P_{α}, we define $f(\mathfrak{a})$ by $\prod_{\alpha} f(P_{\alpha})^{a_{\alpha}}$. If, moreover, $f(\mathfrak{b})$ is defined for a divisor $\mathfrak{b}$ of U, then, $f(\mathfrak{a}+\mathfrak{b})$ is defined, and we have $f(\mathfrak{a}+\mathfrak{b}) = f(\mathfrak{a}) \cdot f(\mathfrak{b})$. These being remarked, we shall prove the following lemma:

LEMMA 4. *Let f and h be two functions on U such that their divisors in the sense of valuation theory coincide, respectively, with (f) and (h). If, moreover, either $f((h))$ or $h((f))$ is defined, so is the other, and they are equal.*

Proof. We first note that nothing is changed by passing to the non-singular model of U. Therefore, we can assume from the beginning that U is non-singular. Also, the first assertion of the lemma is trivial. Let K be a field over which U, f and h are defined; let t be a variable over K. We shall denote by Γ_f the graph of f in the product $U \times D$; similarly for Γ_h. Here, D is, of course, a projective straight line. We define $(h)_t$ by $\Gamma_h \cdot (U \times (t)) = (h)_t \times (t)$, and then $(h)_t$ is prime rational over $K(t)$. Hence, $f((h)_t)$ is rational over $K(t)$, and we can find a function θ on D, which is defined over K, such that $f((h)_t) = \theta(t)$. Then, the divisor of θ is given by $(\theta) = \mathrm{pr}_2[\Gamma_h \cdot ((f) \times D)]$ [5, p. 231]. More explicitly, if we write (f) in the form $\sum_{\alpha} a_{\alpha} \cdot P_{\alpha}$, we get $(\theta) = \sum_{\alpha} a_{\alpha} \cdot h(P_{\alpha})$. Therefore, $\theta(t)$ is of the form $c \cdot \prod_{\alpha} (t - h(P_{\alpha}))^{a_{\alpha}}$ with a certain element c in K, not equal to 0. If we specialize t to 0 and ∞, we get $f((h)_0) = c \cdot h((f))$ and $f((h)_{\infty}) = c$, hence $f((h)) = h((f))$. q. e. d.

Now, let $\mathfrak{a}$ and $\mathfrak{b}$ be two divisors on the non-singular curve C such that $n \cdot \mathfrak{a}$ and $n \cdot \mathfrak{b}$ are, respectively, divisors of functions f and h on C. Then, if either $f(\mathfrak{b})$ or $h(\mathfrak{a})$ is defined, so is the other, and the expression

$$(\mathfrak{a}, \mathfrak{b}; n) = h(\mathfrak{a}) \cdot f(\mathfrak{b})^{-1}$$

is an n-th root of unity by Lemma 4. Furthermore, the same lemma implies that $(\mathfrak{a}, \mathfrak{b}; n)$ depends only on the linear equivalence classes of $\mathfrak{a}$ and $\mathfrak{b}$ on C. Since any linear equivalence class contains a divisor whose components are all generic over a given field, $(\mathfrak{a}, \mathfrak{b}; n)$ induces a multiplication of the group of divisor classes of order n on C with itself.

On the other hand, let ϕ be a canonical function of C which is defined

over $k(u)$. Let t be an arbitrary point of the Jacobian variety J of C, and let $M_1, \cdots, M_{g-1}$ be independent generic points of C over $k(u,t)$. Then, Θ_t is defined as the locus over $k(u,t)$ of the point $\sum_{i=1}^{g-1} \phi(M_i) + t$, and Θ_0 is denoted simply by Θ. Also, if $\mathfrak{b} = \sum_\beta b_\beta \cdot Q_\beta$ is an arbitrary divisor of C, the point $\sum_\beta b_\beta \cdot \phi(Q_\beta)$ of J is denoted by $S[\phi(\mathfrak{b})]$. Put $t = S[\phi(\mathfrak{b})]$. Then, we have $n \cdot t = 0$ if and only if $n \cdot \mathfrak{b}$ is linearly equivalent to zero on C. In this case, if X is a divisor of J which is linearly equivalent to $\Theta_t - \Theta$, we can find a function ϕ on J such that $(\phi) = n \cdot X$. Moreover, if x is a generic point of J over a field of definition K of ϕ, there exists a function ψ on J, which is defined over K, such that $\phi(n \cdot x) = \psi(x)^n$ [6, p. 150]. Now, let s be another point of J of order n. Then, $\psi(x+s) \cdot \psi(x)^{-1}$ is an n-th root of unity, which is denoted by $e_n(s,t)$:

$$\psi(x+s) = e_n(s,t) \cdot \psi(x).$$

As we can see, $e_n(s,t)$ is independent of the choice of X in the linear equivalence class of $\Theta_t - \Theta$. The two n-th roots of unity so obtained are related as follows:

PROPOSITION (Weil). *Let $\mathfrak{a}$ and $\mathfrak{b}$ be two divisors of C, and put $s = S[\phi(\mathfrak{a})]$ and $t = S[\phi(\mathfrak{b})]$. If both s and t are of order n, and if $(\mathfrak{a}, \mathfrak{b}; n)$ is defined, we have $(\mathfrak{a}, \mathfrak{b}; n) = e_n(s,t)$.*

Proof.[4] Since $n \cdot \mathfrak{b}$ is linearly equivalent to zero on C, we can find a function h on C such that $(h) = n \cdot \mathfrak{b}$. Pick a divisor $\mathfrak{m}_0$ of degree g on C which has no point in common with $\mathfrak{b}$. We take an extension K of $k(u)$ over which $\mathfrak{b}$, $\mathfrak{m}_0$ and s are all rational. Let $M_1, \cdots, M_g$ be independent generic points of C over K, and put $\mathfrak{m}_1 = \sum_{i=1}^{g} M_i$ and $x = S[\phi(\mathfrak{m}_1 - \mathfrak{m}_0)]$. Then, $h(\mathfrak{m}_1 - \mathfrak{m}_0)$ is defined, and it is rational over $K(\mathfrak{m}_1) = K(x)$. Therefore, we can find a function ϕ on J, which is defined over K, such that $\phi(x) = h(\mathfrak{m}_1 - \mathfrak{m}_0)$. If $\sum_\beta b_\beta \cdot Q_\beta$ is the reduced expression of $\mathfrak{b}$, and if we put $X = \sum_\beta b_\beta \cdot \Theta_{\phi(Q_\beta)}$, we have $(\phi) = n \cdot X$ (cf. [6], pp. 115-116). Moreover, X is linearly equivalent to $\Theta_t - \Theta$ [6, pp. 105-106]. On the other hand, by Riemann-Roch theorem the linear equivalence class of $\mathfrak{m}_0 + n \cdot (\mathfrak{m}_1 - \mathfrak{m}_0)$ contains at least one positive divisor, say $\mathfrak{m}_n$. Since $S[\phi(\mathfrak{m}_n - \mathfrak{m}_0)] = n \cdot x$ is a generic point of J over K, we conclude that the components of $\mathfrak{m}_n$ are

[4] This proof was communicated to us by Professor André Weil on February 6, 1956. The proposition is interesting, and it simplifies the proof of Theorem 4. In fact, our original proof was direct, and it was much longer than the present one.

g independent generic points of C over K. In particular, $\mathfrak{m}_n$ is unique, and hence it is rational over $K(x)$. Therefore, we can find a function F on C, which is defined over $K(x)$, such that $(F)=\mathfrak{m}_n-\mathfrak{m}_0-n\cdot(\mathfrak{m}_1-\mathfrak{m}_0)$. Let $(\mathfrak{m}_1', F')$ be the transform of $(\mathfrak{m}_1, F)$ by the automorphism of $K(x)$ over $K(n\cdot x)$ mapping x to $x+s$. Then, we have $S[\phi(\mathfrak{m}_1'-\mathfrak{m}_0)]=x+s$, i.e., $S[\phi(\mathfrak{m}_1'-\mathfrak{m}_1)]=s$. In other words, $\mathfrak{m}_1'-\mathfrak{m}_1$ is linearly equivalent to $\mathfrak{a}$. Moreover, we have $(F')=\mathfrak{m}_n-\mathfrak{m}_0-n\cdot(\mathfrak{m}_1'-\mathfrak{m}_0)$. Therefore, if we define a function f on C by $f=F\cdot[F']^{-1}$, it is defined over $K(x)$, and we have $(f)=n\cdot(\mathfrak{m}_1'-\mathfrak{m}_1)$. Furthermore, $\phi(x)=h(\mathfrak{m}_1-\mathfrak{m}_0)$ is transformed into $\phi(x+s)=h(\mathfrak{m}_1'-\mathfrak{m}_0)$. Similarly, if we consider the generic specialization $(\mathfrak{m}_1, x)\rightarrow(\mathfrak{m}_n, n\cdot x)$ over K, we get $\phi(n\cdot x)=h(\mathfrak{m}_n-\mathfrak{m}_0)$. Now, $h(\mathfrak{m}_n-\mathfrak{m}_0)$ can be written as $h(n\cdot(\mathfrak{m}_1-\mathfrak{m}_0))\cdot h((F))$, and this is equal to $(\phi(x)\cdot F(\mathfrak{b}))^n$ by Lemma 4. Since $F(\mathfrak{b})$ is an element of $K(x)$, there exists a function ψ on J, which is defined over K, such that $\psi(x)=\phi(x)\cdot F(\mathfrak{b})$. Then, $\psi(x+s)\cdot\psi(x)^{-1}$ can be written as $(\phi(x+s)\cdot\phi(x)^{-1})\cdot(F'(\mathfrak{b})\cdot F(\mathfrak{b})^{-1})$, and this is equal to $h(\mathfrak{m}_1'-\mathfrak{m}_1)\cdot f(\mathfrak{b})^{-1}$. However, $\psi(x+s)\cdot\psi(x)^{-1}$ is nothing but $e_n(s.t)$, while $h(\mathfrak{m}_1'-\mathfrak{m}_1)\cdot f(\mathfrak{b})^{-1}$ is equal to $(\mathfrak{m}_1'-\mathfrak{m}_1, \mathfrak{b}; n)$ $=(\mathfrak{a}, \mathfrak{b}; n)$. q. e. d.

With these preparations, the proof of the following theorem is a matter of technique:

THEOREM 4. *Let s and t be, respectively, invariant point and vanishing point of order n along I. Then, we have $e_n(s,t)=1$.*

Proof. As in the proof of Theorem 3, we can find a positive divisor $\mathfrak{r}$ of C, which is rational over $k(u)$, such that its specialization $\mathfrak{r}'$ over the specialization $u\rightarrow a$ with reference to k is free from Q, and, moreover, such that the complete linear system $\mathfrak{L}_m-\mathfrak{r}$ is of degree nd. Pick two positive divisors $\mathfrak{m}_0$ and $\mathfrak{n}_0$ of C such that $n\cdot\mathfrak{m}_0$ and $n\cdot\mathfrak{n}_0$ both belong to $\mathfrak{L}_m-\mathfrak{r}$. We can assume that $\mathfrak{m}_0$ and $\mathfrak{n}_0$ are algebraic over $k(u)$, and such that the specializations $\mathfrak{m}_0'$ and $\mathfrak{n}_0'$ of $I(\mathfrak{m}_0)$ and $I(\mathfrak{n}_0)$ at the center of $k[[u-a]]$ have no point in common with $\mathfrak{r}'+Q$, and have no point in common with each other. Pick another two positive divisors $\mathfrak{m}$ and $\mathfrak{n}$ of C of degree d such that $S[\phi(\mathfrak{m}-\mathfrak{m}_0)]=s$ and $S[\phi(\mathfrak{n}-\mathfrak{n}_0)]=t$. Again, we can assume that $\mathfrak{m}$ and $\mathfrak{n}$ are algebraic over $k(u)$, and such that the specializations $\mathfrak{m}'$ and $\mathfrak{n}'$ of $I(\mathfrak{m})$ and $I(\mathfrak{n})$ at the center of $k[[u-a]]$ have no point in common with $\mathfrak{m}_0'+\mathfrak{n}_0'+\mathfrak{r}'+Q$, and have no point in common with each other. Finally, we can find four hypersurfaces F_0, F, H_0 and H in the ambient space of C, all algebraic over $k(u)$, such that they intersect C, respectively, at $n\cdot\mathfrak{m}_0+\mathfrak{r}$, $n\cdot\mathfrak{m}+\mathfrak{r}$, $n\cdot\mathfrak{n}_0+\mathfrak{r}$ and $n\cdot\mathfrak{n}+\mathfrak{r}$. We can assume that the specializations F_0', F', H_0' and H' of $I(F_0)$, $I(F)$, $I(H_0)$ and $I(H)$

at the center of $k[[u-a]]$ do not contain C'. Let $F_0(X)$, $F(X)$, $H_0(X)$ and $H(X)$ be homogeneous polynomials of order m, respectively associated with F_0, F, H_0 and H; similarly for $F_0'(X)$, $F'(X)$, $H_0'(X)$ and $H'(X)$. Then, the ratios $F(X)\cdot F_0(X)^{-1}$ and $H(X)\cdot H_0(X)^{-1}$ define numerical functions f and h on C. Similarly, the ratios $F'(X)\cdot F_0'(X)^{-1}$ and $H'(X)\cdot H_0'(X)^{-1}$ define functions f' and h' on C'. Here, as we can see, we have $(f)=n\cdot(\mathfrak{m}-\mathfrak{m}_0)$, $(h)=n\cdot(\mathfrak{n}-\mathfrak{n}_0)$, $(f')=n\cdot(\mathfrak{m}'-\mathfrak{m}_0')$ and $(h')=n\cdot(\mathfrak{n}'-\mathfrak{n}_0')$. Also, by the construction (f') and (h') are the divisors of f' and h' in the sense of valuation theory. Now, by the previous proposition $h(\mathfrak{m}-\mathfrak{m}_0)\cdot f(\mathfrak{n}-\mathfrak{n}_0)^{-1}$ is equal to $e_n(s,t)$. However, $h'(\mathfrak{m}'-\mathfrak{m}_0')$ and $f'(\mathfrak{n}'-\mathfrak{n}_0')$ are the unique specialization of $h(\mathfrak{m}-\mathfrak{m}_0)$ and $f(\mathfrak{n}-\mathfrak{n}_0)$ over the specialization

$$(u,F_0,F,H_0,H)\rightarrow(a,F_0',F',H_0',H')$$

with reference to k. Since $e_n(s,t)$ is an n-th root of unity, it is invariant under any specialization over k. Therefore, we get

$$e_n(s,t)=h'(\mathfrak{m}'-\mathfrak{m}_0')\cdot f'(\mathfrak{n}'-\mathfrak{n}_0')^{-1}.$$

Since t is a vanishing point along I, there exists a function h'' on C' such that $\mathfrak{n}'-\mathfrak{n}_0'$ is the divisor of h'' in the sense of valuation theory. This does not mean, of course, that h'' is defined by a ratio of homogeneous polynomials which do not vanish at Q. Any way, we get $h'=c\cdot(h'')^n$ with a certain constant c, and hence

$$h'(\mathfrak{m}'-\mathfrak{m}_0')\cdot f'(\mathfrak{n}'-\mathfrak{n}_0')^{-1}=h''(n\cdot(\mathfrak{m}'-\mathfrak{m}_0'))\cdot f'(\mathfrak{n}'-\mathfrak{n}_0')^{-1},$$

and this is equal to 1 by Lemma 4. q. e. d.

We recall, here, that $e_n(s,t)$ is a skew-symmetric bilinear form on the group of points of order n on J [6, p. 153]. The Theorem 4 asserts that the groups of invariant points and vanishing points of order n along I are the groups of annihilators of each other. In particular, each one of them determines the other uniquely.

5. Passage to inductive limits. We shall now consider the limit $n\rightarrow\infty$. Here, n runs over positive integers which are not divisible by the characteristic p, and which "finally" contain powers of any prime l different from p. More conveniently, we consider the limits $l^\nu\rightarrow\infty$ for all l different from p. We fix one such l in the following.

Let $\mathfrak{w}_\nu$ be the additive group of points of J of order l^ν, and let $\mathfrak{w}$ be the union of $\mathfrak{w}_0,\mathfrak{w}_1,\cdots$. If we adjoin all elements of $\mathfrak{w}$ to $k(u)$, we get an algebraic extension K of $k(u)$. Similarly, an algebraic extension $k((u-a))(I(\mathfrak{w}))$ of $k((u-a))$ is defined, and it is the compositum of $I(K)$ and $k((u-a))$. Let $\mathfrak{G}$ and $\bar{\mathfrak{g}}$ be the Galois groups of K over $k(u)$

and $k((u-a))(I(\mathfrak{w}))$ over $k((u-a))$, respectively, with Krull's topologies. If $\bar{\sigma}$ is in $\bar{\mathfrak{g}}$, then $\sigma = I^{-1}\cdot\bar{\sigma}\cdot I$ is in $\mathfrak{G}$. Moreover, the correspondence $\bar{\sigma}\rightarrow\sigma$ is a continuous isomorphism of $\bar{\mathfrak{g}}$ into $\mathfrak{G}$, and we get a compact image group $\mathfrak{g}$. The group $\mathfrak{g}$ is generally called the *decomposition group* of K over $k(u)$ along I. Here, since k is algebraically closed, $\mathfrak{g}$ coincides with the *inertia group* of K over $k(u)$ along I. Since the group law in J is defined over $k(u)$, the group $\mathfrak{G}$, hence in particular the group $\mathfrak{g}$, operates on $\mathfrak{w}$. On the other hand, if we denote by Z_l the ring of l-adic integers, and by Q_l its quotient-field, $\mathfrak{w}$ is isomorphic to a $2g$-fold direct product of the additive group Q_l modulo 1. This isomorphism is called an *l-adic coordinate system* in the group $\mathfrak{w}$. With reference to a fixed l-adic coordinate system, $\mathfrak{G}$ is represented by a group of matrices of degree $2g$ with coefficients in Z_l. We call this matric group the *abstract monodromy group of the fibre system* $\{u\times J\}$ *at* l, while the *abstract local monodromy group* of the fibre system along I at l is defined by using $\mathfrak{g}$ instead of $\mathfrak{G}$. At present, we are interested in the structure of abstract local monodromy groups. We shall prove the following theorem:

THEOREM 5. *We can choose an l-adic coordinate system in $\mathfrak{w}$ such that any element σ of the inertia group $\mathfrak{g}$ is represented by a matrix of the form*[5]

$$\begin{bmatrix} \boxed{\begin{matrix} 1 & c \\ 0 & 1 \end{matrix}} & \\ & \boxed{1_{2g-2}} \end{bmatrix}.$$

Here, the correspondence $\sigma\rightarrow c$ is a continuous isomorphism of $\mathfrak{g}$ into the additive group Z_l.

Proof. We fix, once for all, an isomorphism $\lg$ of the multiplicative group of l^{ν}-th roots of unity for $\nu=0,1,\cdots$ onto the additive group Q_l modulo 1. Also, let $\mathfrak{v}_\nu$ be the group of vanishing points of order l^ν along I, and let $\mathfrak{v}$ be the union of $\mathfrak{v}_0,\mathfrak{v}_1,\cdots$. Then, we can find an isomorphism i of $\mathfrak{v}$ onto Q_l modulo 1. According to Lemma 2, we can extend the isomorphism i to an isomorphism of the group of invariant points of order l^ν for $\nu=0,1,\cdots$ along I onto a $(2g-1)$-fold direct product of Q_l modulo 1. Here, we agree to take the image of $\mathfrak{v}$ as the first factor of the product. Also, we keep the notation i for the extension. Now, let ν be an arbitrary non-negative integer, and let t_ν be an element of $\mathfrak{v}_\nu$ such that $i(t_\nu)\equiv(l^{-\nu}0\cdots 0)$ mod. 1. Then, we have $l\cdot t_{\nu+1}=t_\nu$ for $\nu=0,1,\cdots$. If we denote $e_n(*,*)$,

[5] The symbol 1_{2g-2} denotes a unit matrix of degree $2g-2$: two blanks are zero matrices.

temporarily, by $e_\nu(*,*)$ for $n=l^\nu$, we can find an element, say s_ν, of $\mathfrak{w}_\nu$ such that $1g(e_\nu(s_\nu,t_\nu))\equiv l^{-\nu}$ mod. 1. If we impose the further condition that $e_\nu(s_\nu,s)=1$ for all invariant points s of order l^ν along I satisfying $i(s)\equiv(0*\cdots*)$ mod. 1, then, s_ν is unique modulo $\mathfrak{v}_\nu$. In fact, if s_ν' is another element with similar properties, $s_\nu'-s_\nu$ is annihilated by all invariant points of order l^ν along I. Hence, by Theorem 4 this is an element of $\mathfrak{v}_\nu$. Since the converse is also true, the number or distinct s_ν is equal to l^ν. Since $l\cdot s_{\nu+1}$ is one of the s_ν, we can find a sequence $s_0,s_1,\cdots$ satisfying $l\cdot s_{\nu+1}=s_\nu$ for $\nu=0,1,\cdots$. Now, let s be an arbitrary element of $\mathfrak{w}$, say in $\mathfrak{w}_\nu$; put $1g(e_\nu(s,t_\nu))\equiv n\cdot l^{-\nu}$ mod. 1 with some integer n. Then, by the same reason as above, $s-n\cdot s_\nu$ is an invariant point of order l^ν along I. Hence, we can extend the isomorphism i to an l-adic coordinate system θ in $\mathfrak{w}$ by

$$\theta(s)\equiv(n\cdot l^{-\nu}\ i(s-n\cdot s_\nu))\ \text{mod.}\ 1.$$

In fact, $\theta(s)$ is independent of the choice of ν, and θ gives an isomorphism of $\mathfrak{w}$ onto the $2g$-fold direct product of Q_l modulo 1. Finally, let σ be an arbitrary element of $\mathfrak{g}$, and let ν be a non-negative integer. Then, by Theorem 3 we can find an integer c_ν such that $\sigma(s_\nu)=s_\nu+c_\nu\cdot t_\nu$. Since we have $l\cdot s_{\nu+1}=s_\nu$ and $l\cdot t_{\nu+1}=t_\nu$, we have $c_{\nu+1}\cdot t_\nu=c_\nu\cdot t_\nu$, i.e., $c_{\nu+1}\equiv c_\nu$ mod. l^ν. Therefore, the sequence $c_0,c_1,\cdots$ has a limit, say c, in Z_l such that $c\equiv c_\nu$ mod. l^ν for $\nu=0,1,\cdots$. On the other hand, since s_ν generates $\mathfrak{w}_\nu$ modulo invariant points of order l^ν along I, we conclude that $k((u-a))(I(\mathfrak{w}_\nu))$ coincides with $k((u-a))(I(s_\nu))$. Therefore, the correspondence $\sigma\to c$ gives a continuous isomorphism of $\mathfrak{g}$ into Z_l. Let $M(\sigma)$ be the matrix in the theorem with the above defined c as its $(1,2)$-coefficient. Then, for any element s in $\mathfrak{w}$, we have $\theta(\sigma(s))\equiv\theta(s)\cdot M(\sigma)$ mod. 1. q. e. d.

As we know in general, there exists a skew-symmetric matrix E of degree $2g$ with coefficients in Z_l such that $1g(e_\nu(s,t))\equiv l^\nu\cdot\theta(s)\cdot E\cdot{}^t\theta(t)$ mod. 1 for s and t in $\mathfrak{w}_\nu$ for $\nu=0,1,\cdots$. We know also that E^{-1} exists, and has coefficients in Z_l [6, p. 156]. Now, in the present case, the proof of Theorem 5 shows that E has the following special form:

$$E=\left[\begin{array}{c|c}\begin{array}{cc}0&1\\-1&0\end{array}&\\\hline&E'\end{array}\right].$$

Here, of course, E' is a similar matrix as E.

We shall discuss some supplements of Theorem 5. We first note that the image of $\mathfrak{g}$ in Z_l is compact, hence it is a closed additive subgroup of Z_l.

However, any closed additive subgroup of Z_l is an ideal of the ring Z_l. Therefore, the image in question is of the form $l^\rho \cdot Z_l$ with $0 \leqq \rho \leqq \infty$. Here, we consider l^ρ as 0 for $\rho = \infty$. We shall show that ρ is an *invariant of the fibre system* in the following sense:

SUPPLEMENT 1. *The number ρ depends neither on the isomorphism I nor on the l-adic coordinate system θ.*

In fact, let θ' be a coordinate system in $\mathfrak{w}$ which gives rise to a similar representation of $\mathfrak{g}$ as the one stated in Theorem 5. Let σ be an element of $\mathfrak{g}$, and let $M'(\sigma)$ be the corresponding matrix with respect to the new system. Then, we can find a non-singular matrix L of degree $2g$ such that $M'(\sigma) = L^{-1} \cdot M(\sigma) \cdot L$. Here, both L and L^{-1} have their coefficients in Z_l. Therefore, if c and c' are the $(1,2)$-coefficients of $M(\sigma)$ and $M'(\sigma)$, respectively, we can verify that c and c' differ only by a unit of Z_l. Also the change of I amounts to replace $\mathfrak{g}$ by a conjugate in $\mathfrak{G}$, and the argument is similar.

We note that the invariant ρ does not depend, by Lemma 1, on the choice of the group law of J. Now, in the case of $p = 0$, we can show that $\rho = 0$ for all prime l (cf. [3]). In the modular case, we are satisfied, temporarily, with stating the following remark:

SUPPLEMENT 2. *There exists a linear pencil on the surface V such that the invariant ρ of the associated fibre system of Jacobian varieties along any degenerate fibre is always finite, and also such that $\rho = 0$ for almost all prime l.*

Since our proof of this supplement is not local, we shall reserve it for our global theory.

THE JOHNS HOPKINS UNIVERSITY.

REFERENCES.

[1] W. L. Chow, "The Jacobian variety of an algebraic curve," *American Journal of Mathematics*, vol. 76 (1954), pp. 453-476.

[2] J. Igusa, "Fibre systems of Jacobian varieties," *ibid.*, vol. 78 (1956), pp. 171-199.

[3] ———, "Monodromy groups connected with algebraic surfaces," *to appear.*

[4] M. Rosenlicht, "Generalized Jacobian varieties," *Annals of Mathematics*, vol. 59 (1954), pp. 505-530.

[5] A. Weil, *Foundations of algebraic geometry*, New York (1946).

[6] ———, *Variétés Abéliennes et courbes algébriques*, Paris (1948).

[7] O. Zariski, *Algebraic surfaces*, New York (1948).

[8] ———, "Complete linear systems on normal varieties and a generalization of a lemma of Enriques-Severi," *Annals of Mathematics*, vol. 55 (1952), pp. 552-592.

SIMULTANEOUS RESOLUTION FOR ALGEBRAIC SURFACES.*

By Shreeram Abhyankar.

In Memory of my Late Friend Professor Irvin S. Cohen (1917-1955).

---

1. Introduction. Now that one has the resolution theorem for algebraic surfaces[1] one may naturally ask the following question concerning *simultaneous resolvability*: Given a two dimensional algebraic function field K and a finite algebraic extension K^* of K does there exist a non-singular projective model of K whose K^*-normalization is also non-singular? This and related weaker (global as well as local) questions were raised by us in part B of Section 6 of [4]. The purpose of the present paper is to answer some of these questions. The main results of this paper are equally novel for fields of zero characteristic (in particular for the complex ground field) or for modular fields. From a pair of projective normal models V and V^* respectively of K and K^* such that the transformation between V and V^* is free from fundamental points on both (i.e. V^* is a K^*-normalization of V), instead of requiring that both be non-singular one may require that at least one of them be non-singular. Since requiring V to be non-singular is simply equivalent to the resolution theorem for K it is of no added content; however requiring V^* to be non-singular asks much more about the pair (K, K^*) than the mere resolution theorem for K^* and hence we could call this the question of *weak simultaneous resolvability* for the pair (K, K^*). If we restrict our attention to a zero dimensional valuation of K^* and its K-restriction then we get the problems of *local simultaneous resolution* and *local weak simultaneous resolution* respectively. The results of this paper are quite in conformity with the conjectured statements concerning these questions made in Section 6 of [4]. It is obvious that these simultaneous resolvability questions are related to the problem: Given an involution T on a variety V to remove the fundamental points of these involutions (i.e. the points on V which are fundamental for the rational map of V onto the Chow variety of the involution T). We intend to study this involution problem at some future opportunity.

---

* Received March 15, 1956.

[1] For characteristic zero see Zariski [8], for non-zero characteristic with algebraically closed ground fieelds see Abhyankar [3] and for generalization to perfect ground fields see Abhyankar [4].

Now to describe the results of this paper in greater detail, let K be a two dimensional algebraic function field over ground field k of characteristic p and let K^* be a finite algebraic extension of K. Let v^* be a zero dimensional valuation of K^* and let v be the K-restriction of v^*.

In Section 2 we settle the question of local weak simultaneous resolution in the affirmative; i. e., we prove that if v^* can be uniformized (this condition is certainly satisfied if k is perfect, see [4]), then v^* can be uniformized on a K^*-normalization of a projective normal model of K. This result will then eliminate the use of the Zariski factorization theorem for an antiregular transformation between two surfaces into local quadratic transformation (see Theorem 3 of [2]) from our proof of the local uniformization for algebraic surfaces over algebraically closed modular ground fields [3] and its generalization to perfect ground fields [4]; more explicit references to these points will be made in Section 2. Since generalization of such a precise factorization theorem for higher dimensional varieties is not possible, this elimination of the factorization theorem was essential before attempting uniformization for higher varieties (over modular ground fields). In Section 3 we prove that if k is algebraically closed, $p=0$ and v is rational, then local simultaneous resolution holds, i. e., there exists a projective normal model V of K on which v has a simple center and such that on the K^*-normalization of V the center of v^* is simple. In Section 4 we give an affirmative answer to the (global) weak simultaneous resolvability question in case k is algebraically closed, $p=0$ and K^*/K is galois.

The more novel part of the paper (Sections 7 and 8) dealing with (global) simultaneous resolution is preceded by a few words on the branch locus (Section 5) and some theorems on kummer extensions of unique factorization domains (Section 6). In Section 7, assuming k to be perfect, we prove that either if K^*/K is a quadratic extension and $p\neq 2$ or if K^*/K is a cubic cyclic extension, $p\neq 3$ and k contains primitive cube roots of unity then there exists a *non-singular* projective model V of K whose K^*-normalization is also *non-singular* and the branch locus on V is a non-singular (in general reducible) curve. It is expected that this result should be helpful towards a comparison of the various structural properties and associated invariants of the fields K and K^*. In Section 8, we prove that the answer to the simultaneous resolution question, although positive for quadratic and cubic cyclic extensions, is negative in general. More precisely we prove that if K possesses minimal models (i. e., if K is not the function field of a ruled surface) and if q is an arbitrary prime number with $q\neq p$ and $q>3$ then there exist (lots of) q-cyclic extensions K' of K such that for any non-singular projective model of K the K'-normalization is necessarily singular.

Finally in Section 9 we give partial generalizations to higher dimensional varieties of the results of Sections 2, 3 and 7. These considerations might throw some light on Zariski's proposed method of resolution of singularities for higher varieties [13].

The notations in this paper will be the same as those used in [1], [3] and [4].

2. Uniformization on a derived normal model. In our proof of the local uniformization theorem for algebraic surfaces over algebraically closed modular ground fields ([3], see proof of Theorem 3 in Section 6 on "Cyclic extensions of prime degree $q \neq p$; coming down") and also in its generalization to perfect ground fields ([4], see Theorem 3 of Section 2) we had to deal with the following situation: Given a galois extension K^* of a two dimensional algebraic function field K and a zero dimensional valuation v of K having a unique extension v^* to K^*, knowing that v^* can be uniformized the problem was to uniformize v^* on a K^*-normalization of some projective normal model of K. This was achieved by using the Zariski factorization theorem of antiregular transformations between algebraic surfaces into local quadratic transformations (Theorem 3 of [2]). A second instance where we faced a similar situation in our uniformization proof was as follows (Theorem 1 of Section 4 of [3] and Theorem 2 of Section 2 of [4]): K^* is a finite separable algebraic extension of a two dimensional algebraic function field K, v^* a zero dimensional valuation of K^* and v the K-restriction of v^*, such that for a projective normal model V of K and its K^*-normalization V^* we have $d(R^*:R)=1$ where R^* and R are the quotient rings of the centers of v^* and v respectively on V^* and V. Given that v^* can be uniformized the problem was to uniformize v^* on a K^*-normalization of some projective normal model of K which dominates V.

Both of these are obviously special cases of the weak local simultaneous resolution theorem for algebraic surfaces, i.e., Theorem 1 below. The proof of Theorem 1 in case v is of rational rank greater than one (this is the easy case) is essentially the same as of Theorem 2 of [4] while for the case when v is rational we have to use one new trick. Consequently the only deep property of quadratic transformations of two dimensional regular local domains which is used in the proof of Theorem 1 is Theorem 2 of [2] (which may be called weak local uniformization theorem). Since both the instances of the use of the Zariski factorization theorem in our proof of the local uniformization theorem for algebraic surfaces are special cases of Theorem 1 below, we have now eliminated the use of the factorization theorem. This

is important for the uniformization problem for higher varieties (over modular ground fields), since to generalize the factorization theorem, in any form, to higher varieties is a very difficult problem at all events (even in the classical case). On the other hand the weak local uniformization theorem is true for higher varieties at least over algebraically closed ground fields of characteristic zero (see Section 9). Besides these applications the following theorem on local weak simultaneous is of importance in itself.

THEOREM 1. (*Local weak simultaneous resolution*). *Let K/k be a two dimensional algebraic function field and let K^* be a finite algebraic extension of K. Let v^* be a zero dimensional valuation of K^*/k and let v be the K-restriction of v^*. Assume that v^* can be uniformized. Then v^* can be uniformized on the K^*-normalization of a projective normal model of K/k.* [More precisely if R^* is the quotient ring of a simple center on some projective normal model of K/k then we can find a quadratic transform of R^* along v^* which is a quotient ring on a K^*-normalization of a projective normal model of K/k.]

Proof. (In general we shall follow the proof of Theorem 2 of [4] or rather we shall sharpen that proof). Let (R^*, M^*) be the quotient ring of a simple center of v^* on some projective normal model of K^*/k. Let $S = R^* \cap K$, $Q = M^* \cap K$ and $P = QR^*$. Now if e is an element of K which is integral over S then e is integral over R^*, i.e. e is in R^* and hence e is in $R^* \cap K = S$. Therefore S is integrally closed in K. Hence $K =$ the quotient field of S if and only if tr. deg. (transcendence degree of) $S/k =$ tr. deg. K/k. We proceed to arrange matters so that tr. deg. S/k $=$ tr. deg. K/k. Let q_1, q_2 be a transcendence basis of K/k. Replacing q_i by $1/q_i$ if necessary, we may assume that $q_1, q_2 \in R_v$. Replacing R^* by a quadratic transform (see Definition 3 of [2]) along v^* we may assume that $q_1, q_2 \in R^*$ (Lemma 12 of [2]), i.e. that K is the quotient field of S.

Now fix a non-zero element z in M_v. Again replacing R^* by a quadratic transform along v^* we may assume that (Theorem 2 of [2]):

$$z = x^a y^b d$$

where (x, y) is a basis of M^*, a and b are non-negative integers, d is a unit in R^* and where we have either that $b = 0$ and $a > 0$ or that $v^*(x)$ and $v^*(y)$ are rationally independent. We proceed to show that in either case P is primary for M^*; and that will then, in view of Proposition 1 of [4], complete the proof.

Case 1, $a > 0$ *and* $b = 0$. Now K is the quotient field of $S = R^* \cap K$

3. Simultaneous uniformization of rational valuations. The aim of this section is to prove

THEOREM 2. (*Local simultaneous resolution for rational valuations*). *Let K be a two dimensional algebraic function field over an algebraically closed ground field k of characteristic zero and let v be a zero dimensional rational valuation of K/k. Let K^* be a finite algebraic extension of K and let v^* be an extension of v to K^*. Then there exists a projective normal* (in fact non-singular) *model V of K/k on which the center of v is at a simple point and such that the center of v^* on a K^*-normalization of V is also a simple point.*

Before proving this theorem we must make the ramification theoretic preparation for it. The following proposition on rational groups was conjectured by us and a proof of it was kindly provided by Professor Iwasawa.

PROPOSITION 1. *Let G be a torsion free abelian group of rational rank one and let H be a subgroup of G of finite index. Then G/H is cyclic.*

Proof. We may arrange matters so that G is a subgroup of the additive group R of rational numbers and so that the subgroup Z of integers is contained in H. Let $R^* = R/Z$, $G^* = G/Z$ and $H^* = H/Z$. Then G/H is isomorphic to G^*/H^* and hence it is enough to show that G^*/H^* is cyclic. Now the elements of R^* can be considered to be rational numbers r with $0 \leqq r < 1$, and the addition in R^* then being addition modulo 1. Given integers r and s with $0 \leqq r < s$, we write $s = \prod_{i=1}^{n} p_i^{u_i}$, where $p_1, p_2, \cdots, p_n$ are distinct prime numbers and $u_1, u_2, \cdots, u_n$ are positive integers. There exist, by partial fraction theory, unique integers r_i with $0 \leqq r_i < p_i^{u_i}$ such that $r/s = \sum_{i=1}^{n} r_i/p_i^{u_i}$. Thus R^* has the direct sum representation: $R^* = \oplus_{p_i \in P} S(p_i)$ where P is the set of all prime numbers and

$$S(p_i) = S_\infty(p_i) = \bigcup_{j=1}^{\infty} S_j(p_i)$$

where $S_j(p_i)$ is the subgroup of elements t of R^* for which $p_i^j t$ is an integer. Now let t be any given non-zero element of G^*. We can write $t = r/s$ where r and s are coprime integers such that $0 < r < s$. Let $s = \prod_{i=1}^{n} p_i^{u_i}$ where $p_1, p_2, \cdots, p_n$ are distinct prime numbers and $u_1, u_2, \cdots, u_n$ are positive integers. Then there exist unique integers r_i with $0 < r_i < p_i^{u_i}$ such that

$t=\sum_{i=1}^{n} r_i/p_i^{u_i}$. If $n=1$ then obviously $t \in G^* \cap S(p_1)$. Now suppose $n>1$. Then we can find integers a and b such that

$$bp_1^{u_1}+a\prod_{i=2}^{n} p_i^{u_i}=1;\ \text{so that}\ (a\prod_{i=2}^{n} p_i^{u_i})r_1=r_1+cp_1^{u_1},$$

where c is the integer $(-br_1)$. Hence

$$(a\prod_{i=2}^{n} p_i^{u_i})t=r_1p_1^{-u_1}+c+\sum_{i=2}^{n} ar_id_i,$$

where d_i is the integer $(\prod_{j=2}^{n} p_j^{u_j})p_i^{-u_i}$ so that $r_1p_1^{-u_1} \in G^* \cap S(p_1)$. Similarly $r_ip_i^{-u_i} \in G^* \cap S(p_i)$ for $i=2,3,\cdots,n$. Therefore $G^*=\oplus_{p_i \in P} G^* \cap S(p_i)$ and similarly $H^*=\oplus_{p_i \in P} H^* \cap S(p_i)$. Now $G^* \cap S(p_i)=S_{v_i}(p_i)$ where v_i is either a non-negative integer or the symbol ∞ and similarly $H^* \cap S(p_i)=S_{w_i}(p_i)$. We have that $w_i \leqq v_i$ and $G^*/H^*=\oplus_{p_i \in P}[S_{v_i}(p_i)/S_{w_i}(p_i)]$. Since G^*/H^* is finite we must have $v_i=w_i$ except for a finite number of subscripts $e_1, e_2, \cdots, e_m$ for which $v_{e_q}<w_{e_q}<\infty$ (for $q=1,2,\cdots,m$). Since $S_{v_i}(p_i)/S_{w_i}(p_i)$ is cyclic of order $p_i^{v_i-w_i}$ for $i=e_1,e_2,\cdots,e_m$; we conclude that G^*/H^* is cyclic of order $\prod_{j=e_1,e_2,\cdots,e_m} p_j^{(v_j-w_j)}$.

PROPOSITION 2. *Let K be an algebraic function field over an algebraically closed ground field k of characteristic zero and let K^* be a galois extension of K. Let v^* be a zero dimensional rational valuation of K^*/k and let v be the K-restriction of v^*. Then the splitting group of v^* over v is cyclic.*

Proof. The proof follows from Proposition 1 above and Satz 3 of Krull [5], in view of the fact that the residue field of v is the algebraically closed field k of characteristic zero.

PROPOSITION 3. *Let K be a two dimensional algebraic function field over an algebraically closed ground field k and let K^* be a galois extension of K. Let (R,M) be the quotient ring of a simple point on some projective model of K/k. Let v be a real zero dimensional valuation of K/k having center M in R and let v^* be an extension of v to K^*. Let R_n be the n-th quadratic transform* (Definition 3 of [2]) *of R along v and let (R_n^*, M_n^*) be the local ring in K^* lying above R_n such that v^* has center M_n^* in R_n^*. Then there exists an integer m such that for all $n \geqq m$ the splitting field of R_n^* over R_n is the same as the splitting field of v^* over v.*

The proof of this proposition is contained in the proof of Proposition 4

of Section 4 of [3]. That proof was based among other things on the fact that $\bigcup_{n=1}^{\infty} R_n = R_v$ (Lemma 10, Section 3 of [3]). Another proof based on the weak local uniformization theorem (Proposition 3, Section 3 of [3]) will be given Section 9 (Proposition 3A) and this proof will be applicable to higher dimensional varieties over ground fields of characteristic zero which are algebraically closed. In [3], the weak local uniformization was deduced from the theorem: $\bigcup_{n=1}^{\infty} R_n = R_v$. However, in Section 9 the weak uniformization theorem will be proved for higher varieties (real valuations, characteristic zero) and in particular for surfaces without the use of the theorem: $\bigcup_{n=1}^{\infty} R_n = R_v$. This last theorem has no known generalization to higher varieties.

Proof of Theorem 2. (We shall freely use the results of Section 2 of [3] without making specific references.) By the local uniformization theorem, we can find a projective normal model of K/k on which the center of v is a simple point. Let (R, M) be the quotient ring of this simple point. Let K' be a galois extension of K containing K^* and let v' be an extension of v^* to K'. Let K_s' be the splitting field of v' over v and let K_s^* be the compositum of K^* and K_s'. By Proposition 2, K'/K_s' is cyclic and hence K_s^*/K_s' is also cyclic. Let (R', M') be the local ring in K' lying above R such that v' has center M' in R'. Let R_s', R_s^*, R^* and M_s', M_s^*, M^* be the respective intersections of R' and M' with K_s', K_s^* and K^*. By Proposition 3 we may assume that K_s' is the splitting field of R' over R. Then K_s^* is the splitting field of R' over R^*. Let (x, y) be a basis of M. Then (x, y) is a basis also of M_s'. We can find a primitive element z of K_s^*/K_s' with minimal polynomial $X^n - t$ where t is in R_s'. If $n = 1$ there would be nothing to prove; so assume that $n > 1$. By the weak uniformization theorem (Lemma 10 of [3]), after replacing R_s' by a quadratic transform (S, N) along the K_s'-restriction of v', we may arrange matters so that $t = f^u d$ where (f, g) is a basis of N, u is a non-negative integer and d is a unit in S. Since k is algebraically closed, and (x, y) is a basis of M_s' as well as of M, it is clear that after replacing R by its corresponding quadratic transform along v we may assume that $S = R_s'$.

Suppose, if possible, that u and n have a common factor $m > 1$. Since k is algebraically closed of characteristic zero, $X^n - f^u d$ factors in $k[[f, g]][X]$ into m pairwise coprime fatcors. Since $k[[f, g]]$ is the completion of R_s' we have that R_s' splits in K_s^* into more than one local ring; this is a contradiction. Therefore u is prime to n. Hence by the argument used in the proof of Theorem 2 of Section 5 of [3] we conclude that R_s^* is regular.

Since K_s^* is the splitting field of R' over R^*, the degree of R_s^* over R^* is one and hence R^* is also regular.

4. Resolution on a derived normal model. The purpose of this section is to prove

THEOREM 3. *Let K be a two dimensional algebraic function field over an algebraically closed ground field k of characteristic zero and let K^* be a galois extension of K. Then there exists a projective normal model of K/k whose K^*-normalization is non-singular.*

The proof of this theorem will consist of several remarks.

(1) (This remark is valid for an arbitrary finite algebraic extension K^* of any dimensional algebraic function field K over an arbitrary ground field k). Let V^* be a projective model of K^*/k and let T be the rational transformation from V^* onto the Chow variety V of sets of points of the involution on V^* defined by the subfield K of K^*. Let W^* be a normalization of V^* and let U be the rational transformation from W^* onto the Chow variety W of sets of points of the involution on W^* defined by the subfield K of K^*. Finally let W_1 be a normalization (in K) of W. Then we have that

(a) T has no fundamental point on V^* if and only if U has no fundamental points on W^*.

(b) U has no fundamental points on W^* if and only if W^* is a K^*-normaliaztion of some projective normal model of K/k. In fact if U has no fundamental points on W^* then W^* is a K^*-normalization of W_1.

(2) Let V be a projective normal model of K/k and let $Q = (P_1, P_2, \cdots, P_n)$ be a finite set of (distinct) points of V. Let $V^{(1)}$ be the surface obtained by applying a local quadratic transformation to V centered at P_1 and followed by normalization. Let $P_2^{(1)}$ be the point on $V^{(1)}$ which corresponds to P_2. Let $V^{(2)}$ be the surface obtained by applying a local quadratic transformation to $V^{(1)}$ centered at $P_2^{(1)}$ and followed by normalization. And so on until we obtain a normal model $V^{(n)}$ of K/k. It is clear $V^{(n)}$ depends (up to a well defined biregular transformation) only on the set Q and not on the order in which the points P_i are taken. We shall say that $V^{(n)}$ is obtained from V by applying a quadratic transformation centered at the set of points Q and followed by normalization.

(3) Let V^* be a K^*-normalization of some projective normal model of K/k. Then the singular points of V^* (which are finite in number) arrange

themselves in complete K-conjugate sets of points (since K^*/K is galois). Let $P_1, P_2, \cdots, P_n$ be such a complete K-conjugate set of singular points of V^* and let W^* be the surface obtained from V^* by applying a quadratic transformation centered at $(P_1, P_2, \cdots, P_n)$ and followed by normalization. Then it follows by (1) that W^* is K^*-normalization of a projective normal model of K/k.

(4) Let again V^* be a K^*-normalization of some projective normal model of K/k and let $P_1, P_2, \cdots, P_N$ be the set of all the singularities of V^*. Let V_1^* be the surface obtained from V^* by applying a quadratic transformation centered at $(P_1, P_2, \cdots, P_N)$ and followed by normalization. Then (2) and (3) tells us that V_1^* is a K^*-normalization of some projective normal model of K/k.

(5) Let the notation be as in (4). Let $g(P_i)$ be the character of P_i as defined in Part VI of Zariski [8]. Let $h = \max(g(P_1), g(P_2), \cdots, g(P_N))$. Let V_2^* be the surface obtained from V_1^* as V_1^* was obtained from V^*. So on until we have obtained a projective normal model V_h^* of K^*/k which [by (4)] is a K^*-normalization of some projective normal model of K/k. In view of (2), it follows from the proof given by Zariski in Part VI of [8] that V_h^* is non-singular.

5. Preliminaries on the branch locus. In Lemmas 1 and 2 of [1] we had collected well known facts concerning the notion of derived normal model of an algebraic variety in a finite algebraic extension of its function field due to Zariski; there we unnecessarily restricted the extension to be separable and the ground field to be algebraically closed. For completeness we restate

LEMMA 1. *Let K be an r-dimensional algebraic function field and K^* a finite algebraic extension of K. Let V and V^* be normal projective models respectively of K and K^* such that the rational transformation from V^* onto V and its inverse map T^{-1} are both free from fundamental points. For a fixed irreducible subvariety W of V let $W_1^*, W_2^*, \cdots, W_s^*$ be the irreducible subvarieties of V^* corresponding to W. Let R and R_i^* be the quotient rings respectively of W and W_i^* on V and V^*. Then $R_1^*, R_2^*, \cdots, R_s^*$ are exactly the local rings in K^* lying above R.*

Proof. Let A be an affine coordinate ring of V contained in R. Let V' be a projective model of K^* having the integral closure A^* of A in K^* as an affine coordinate ring. Then to W there correspond on V' a finite

number of irreducible subvarieties $W_1', W_2', \cdots, W_t'$; V' is normal at $W_1', W_2', \cdots, W_t'$ and their quotient rings on V' are exactly the local rings in K^* lying above R. Furthermore it is clear that the birational transformation from V' *to* V^* is biregular at $W_1', W_2', \cdots, W_t'$ and maps these onto $W_1^*, W_2^*, \cdots, W_s^*$. Hence $s = t$ and after a suitable rearrangement we have $R_i = Q(W_i^*. V^*) = Q(W_i', V')$ for $i = 1, 2, \cdots, t$.

LEMMA 2. *Given K, K^* and V as in Lemma* 1, *V^* exists as in Lemma* 1 *and is unique up to a biregular transformation.*

Proof. Unicity follows from Lemma 1. Existence is proved on pages 68-70 of [12] or also in [6].

We call V^* a derived normal model of V in K^* or a K^*-normalization of V. By the *branch locus* $B(T)$ on V of the rational transformation T from V^* onto V we shall mean the set of all (*algebraic*) points P of V such that P is on some irreducible subvariety of V which is a branch subvariety for T.

LEMMA 3. *An irreducible subvariety W of V is a branch subvariety for T if and only if $W \subset B(T)$. Furthermore, $B(T) = V$ or $B(T)$ is a proper subvariety of V according as K^*/K is inseparable or separable.*

Proof. The proof is the same as that of Lemma 3 of [1] except that in the alternative proof given there in terms of symmetric product, any reference to "rational points" should be replaced by "algebraic points."

Remark 3. Let K/k be an r-dimensional algebraic function field and K^* a finite separable algebraic extension of K. Let V be a non-singular projective model of K/k and let V^* be K^*-normalization of V. Let D be the branch locus on V. If k is algebraically closed, then Zariski's theorem (Theorem 1 of [1]) tells us that D is pure $r-1$ dimensional. The proof of this theorem is based on Zariski's Jacobian Criterion and hence it is expected that this theorem could be generalized to the case when k is perfect, but if k is imperfect then one would have to replace the Jacobian Criterion by Zariski's Mixed Jacobian Criterion and that would *a priori* necessitate replacing the Jacobian determinant J in the proof of Theorem 1 of [1] by the various $m \times m$ subdeterminants of a certain $m \times M$ Jacobian matrix with $M > m$; hence one cannot predict whether the purity of the branch locus holds true for imperfect ground fields and this point should be looked into. However, it will be a consequence of Theorem 4 proved in the next section that if K^* is cyclic extension of prime degree q which is different from the characteristic of the arbitrary ground field k and if k contains a

primitive q-th root then D is pure $r-1$ dimensional (see Theorem 7 of the next section).

6. Kummer extensions of unique factorization domains.

The following is the main theorem of this section.

THEOREM 4. *Let R be a unique factorization domain with quotient field K of characteristic p (which may or may not be zero). Let K^* be a cyclic extension of K of prime degree $q \neq p$, $(q > 1)$. Assume that K contains a primitive q-th root of unity. Let R^* be the integral closure of R in K^*. Let z be a primitive element of K^*/K for which $z^q = x \in R$ (such elements z do exist). Let $x = dy^q \prod_{i=1}^{s} x_i^{n_i}$ where $y \in R$; $x_1, x_2, \cdots, x_s$ are distinct (i.e. pairwise coprime) irreducible non-units in R, $0 < n_i < q$, and d is a unit in R. Let $jn_i = t_{ij}q + m_{ij}$ with integers t_{ij} and m_{ij} such that $t_{ij} \geqq 0$ and $0 < m_{ij} < q$ for $j = 1, 2, \cdots, q-1$. Let*

$$g_j = z^j y^{-j} \prod_{i=1}^{s} x_i^{-t_{ij}} \qquad \textit{for } j = 1, 2, \cdots, q-1.$$

Then:

(I) *$R^* = R + g_1R + g_2R + \cdots + g_{q-1}R$ where the sum is R-direct; and*[2]

(II) $$D(1, g_1, g_2, \cdots, g_{q-1}) = (-1)^J q^q d^{q-1} (x_1 x_2 \cdots x_s)^{q-1},$$

where $J = (q-1)/2$ if $q > 2$ and $J = 0$ if $q = 2$; and hence

$$D(R^*/R) = \Big(\prod_{i=1}^{s} x_i^{q-1}\Big) R.$$

Proof. Primitive elements z of K^*/K for which $z^q = x \in K$ exist since K contains primitive q-th roots of unity [$z^* \in K^*$ is another primitive element of K^*/K for which $z^{*q} \in K$ if and only if $z^* = uz^i$ with $0 \neq u \in K$ and $0 < i < q$]. Since K is the quotient field of R we may arrange matters so that $x \in R$.

Obviously $K^* = K + g_1K + g_2K + \cdots + g_{q-1}K$, (direct K-sum). Now

$$g_j^q = z^{jq} y^{-jq} \prod_{i=1}^{s} x_i^{-t_{ij}q} = z^{jq} y^{-jq} \Big(\prod_{i=1}^{s} x_i^{-t_{ij}q - m_{ij}}\Big) \Big(\prod_{i=1}^{s} x_i^{m_{ij}}\Big)$$

$$= z^{jq} y^{-jq} \Big(\prod_{i=1}^{s} x_i^{-jn_i}\Big) \Big(\prod_{i=1}^{s} x_i^{m_{ij}}\Big) = \Big(z^q/dy^q \prod_{i=1}^{s} x_i^{n_i}\Big)^j d^j \prod_{i=1}^{s} x_i^{m_{ij}} = d^j \prod_{i=1}^{s} x_i^{m_{ij}} \in R.$$

[2] We are using the following notation: $D(R^*/R)$ = the discriminant ideal of R^* over R = the ideal generated in R by all the K-discriminants $D(u_1, u_2, \cdots, u_q)$ of the various K-bases $(u_1, u_2, \cdots, u_q)$ of K^* consisting of elements in R^*.

Therefore $g_j \in R^*$ for $j=1,2,\cdots,q-1$, and hence $(R+g_1R+g_2R+\cdots+g_{q-1}R) \subset R^*$ where the sum is R-direct. Thus to complete the proof of (I) we have to show that $R^* \subset (R+g_1R+g_2R+\cdots+g_{q-1}R)$.

Let $h=z/y$. Then h is a primitive element of K^*/K with minimal monic polynomial X^q-f where $f=d\prod_{i=1}^{s} x_i^{n_i} \in R$. Since q is prime and since $0<n_i<q$, there exist integers M_i, N_i such that $0<M_i<q$ and $M_i n_i+N_i q=1$. Let $h_i=h^{M_i}x_i^{N_i}$. Then h_i is a primitive element of K^*/K with minimal monic polynomial X^q-f_i with $f_i=x_iF_i$ where F_i is an element of R not divisible by x_i. Let v_i be the real discrete valuation of K which is given by the condition that for r in R we have $v_i(r)=$ the exponent of the highest power of x_i which divides r in R. Then $v_i(h_i^q)=1$ and hence v_i has a unique extension w_i to K^*. We normalize w_i so that $w_i(h_i)=1$. Then $w_i(k)=qv_i(k)$ for any k in K.

Now discriminant $(X^q-f)=(-1)^J q^q f^{q-1}$. Therefore (see pages 79-80 of [7]), given $0\neq a^*\in R^*$ we can find elements a_i in K such that

$$(1) \qquad a^*=a_0+a_1h+a_2h^2+\cdots+a_{q-1}h^{q-1} \text{ and } f^{q-1}a_j\in R.$$

Suppose, if possible, that $w_i(a_jh^j)=w_i(a_kh^k)$ for some $j\neq k$ with $a_j\neq 0\neq a_k$. Then $n_i(j-k)=w_i(h^{j-k})=w_i(a_k/a_j)=qv_i(a_k/a_j)\equiv 0 \pmod{q}$ which is a contradiction since n_i and $(j-k)$ are both prime to q. Therefore no two non-zero terms on the right hand side of (1) are of equal w_i value. Since $a^*\in R^*$, $w_i(a^*)\geqq 0$ and hence $w_i(a_jh^j)\geqq 0$ for $j=1,2,\cdots,q-1$, (if $a_j=0$ then $w_i(a_jh^j)=\infty$). Let $b_j=a_j\prod_{i=1}^{s}x_i^{t_{ij}}$. Then

$$(2) \qquad a^*=b_0+b_1g_1+b_2g_2+\cdots+b_{q-1}g_{q-1} \text{ and } f^{q-1}b_j\in R$$
$$\text{for } j=1,2,\cdots,q-1.$$

Now $w_i(h^j)=(1/q)w_i(h^{qj})=v_i(h^{qj})=v_i(f^j)=t_{ij}q+m_{ij}$. Therefore

$$qv_i(a_j)=w_i(a_j)\geqq -t_{ij}q-m_{ij}, \text{ i.e., } v_i(a_j)\geqq -t_{ij}-(m_{ij}/q).$$

Since $0<(m_{ij}/q)<1$ we must have $v_i(a_j)\geqq -t_{ij}$. Hence $v_i(b_j)=v_i(a_j)+t_{ij}\geqq 0$, i.e., x_i does not occur in the denominator of b_j. This is true for $i=1,2,\cdots,s$. Since $(dx_1^{n_1}x_2^{n_2}\cdots x_s^{n_s})^{q-1}b_j\in R$, we must have that $b_j\in R$ for $j=0,1,\cdots,q-1$. Hence in view of the equation (2) we have that $R^*\subset(R+g_1R+g_2R+\cdots+g_{q-1}R)$ and hence that $R^*=(R+g_1R+g_2R+\cdots+g_{q-1}R)$.

Since $z^j=y^j(\prod_{i=1}^{s}x_i^{t_{ij}})g_j$ and since X^q-x is the minimal monic polynomial of z over K we have that

$$(-1)^J q^q x^{q-1} = \text{Discriminant of } (X^q - x) = D(1, z, z^2, \cdots, z^{q-1})$$
$$= (\prod_{j=1}^{q-1} [y^j \prod_{i=1}^{s} x_i^{t_{ij}}])^2 D(1, g_1, g_2, \cdots, g_{q-1})$$
$$= y^{q(q-1)} (\prod_{j=1}^{q-1} \prod_{i=1}^{s} x_i^{t_{ij}})^2 D(1, g_1, g_2, \cdots, g_{q-1})$$

Therefore $D(1, g_1, g_2, \cdots, g_{q-1}) = (-1)^J q^q d^{q-1} \prod_{i=1}^{s} x_i^{e_i}$, where

$$e_i = (q-1)n_i - 2\sum_{j=1}^{q-1} t_{ij} \geqq 0.$$

Let d_u be a unit in R; y_u an element in R; and $x_{u1}, x_{u2}, \cdots, x_{us_u}$ with $x_{u1} = x_u$ be distinct (i. e. pairwise coprime) irreducible non-units in R. Let $h_u = d_u y_u^q \prod_{i=1}^{s_u} x_{ui}^{n_{ui}}$ where n_{ui} are integers such that $0 < n_{ui} < q$. Then $n_{u1} = 1$. Let t_{uij} and m_{uij} be integers with $t_{uij} \geqq 0$ and $0 < m_{uij} < q$ such that $jn_{ui} = t_{uij} + m_{uij}$ for $j = 1, 2, \cdots, q-1$. Let $g_{uj} = h_u^j y_u^{-j} \prod_{i=1}^{s_u} x_{ui}^{-t_{uij}}$. Replacing z by h_u we coiclude that $R^* = (R + g_{u1}R + g_{u2}R + \cdots + g_{u(q-1)}R)$ and

$$D(1, g_{u1}, g_{u2}, \cdots, g_{u(q-1)}) = (-1)^J q^q d_u^{q-1} \prod_{i=1}^{s_u} x_{ui}^{e_{ui}}$$

where $e_{ui} = (q-1)n_{ui} - 2\sum_{j=1}^{q-1} t_{uij} \geqq 0$. Since $n_{u1} = 1$ we must have $t_{u1j} = 0$ for $j = 1, 2, \cdots, q-1$. Therefore $e_{u1} = (q-1)n_{u1} = q-1$. Since $(1, g_1, g_2, \cdots, g_{q-1})$ and $(1, g_{u1}, g_{u2}, \cdots, g_{u(q-1)})$ are both R-bases of R^* we must have $D(1, g_{u1}, g_{u2}, \cdots, g_{u(q-1)}) = d_u^* D(1, g_1, g_2, \cdots, g_{q-1})$ where d_u^* is a unit in R. Therefore for $u = 1, 2, \cdots, s$, we have that

$$e_u = v_u(D(1, g_1, g_2, \cdots, g_{q-1})) = v_u(D(1, g_{u1}, g_{u2}, \cdots, g_{u(q-1)})) = e_{u1} = q-1.$$

Q. E. D.

COROLLARY. *The above proof yields the arithmetical result:*

$$(q-1)n_i - 2\sum_{j=1}^{q-1} t_{ij} = q-1.$$

A direct proof of this corollary: Since q is prime and since $0 < n_i < q$, it follows that $(m_{i1}, m_{i2}, \cdots, m_{i(q-1)})$ must be a permutation of $(1, 2, \cdots, q-1)$. Hence if we let all the summations Σ be taken over $j = 1, , \cdots, q-1$, we obtain:

$$\Sigma j n_i = \Sigma t_{ij} q + \Sigma m_{ij} = q \Sigma t_{ij} + q(q-1)/2.$$

Therefore

$$q\Sigma t_{ij} = -[q(q-1)/2] + (\Sigma j)n_i = [q(q-1)/2](-1+n_i).$$

Therefore

$$2\Sigma t_{ij} = (q-1)(n_i-1) = (q-1)n_i - (q-1).$$

Therefore

$$(q-1)n_i - 2\Sigma t_{ij} = q-1.$$

THEOREM 5. *In the notation of Theorem 4 assume that R is a local domain with maximal ideal M and that $s \geqq 1$. Then R^* is also a local domain. Let B be a basis of M. Then $(B, g_1, g_2, \cdots, g_{q-1})R^* = M^*$ where by M^* we are denoting the maximal ideal in R^*.*

Proof. We shall use the notation of the proof of Theorem 4. Since v_i has a unique extension to K^*, there exists a unique local domain in K^* lying above R and it must be R^*. Now for $i = 1, 2, \cdots, q-1$ we have that $g_i \in M$ and hence $g_i \in M^*$. Therefore $(B, g_1, g_2, \cdots, g_{q-1})R^* \subset M^*$. Let $m^* \in M^*$. Then $m^* \in R^*$ and hence $m^* = m_0 + m_1 g_1 + m_2 g_2 + \cdots + m_{q-1} g_{q-1}$ with $m_i \in R$. Since $g_i \in M^*$ we have

$$m_0 = m^* - m_1 g_1 - m_2 g_2 - \cdots - m_{q-1} g_{q-1} \in M^* \cap R = BR$$

and hence $m^* \in (B, g_1, g_2, \cdots, g_{q-1})R^*$.

THEOREM 6. *In the notation of Theorem 5 assume that R is noetherian and regular and that $s \geqq 2$. Then R^* is non-regular.*

Proof. We shall use the notation of the proof of Theorem 4. Let $A_0, A_1, \cdots, A_H$ be a minimal basis of M. Suppose, if possible, that there exist elements $a_1, a_2, \cdots, a_H$, $b_1, b_2, \cdots, b_{q-1}$ in R^* such that

$$A_0 = \sum_{i=1}^{H} a_i A_i + \sum_{i=1}^{q-1} b_i g_i. \tag{1}$$

Let a_{ij}, b_{ij} be elements in R such that

$$a_i = a_{i0} + \sum_{j=1}^{q-1} a_{ij} g_j \quad \text{and} \quad b_i = b_{i0} + \sum_{j=1}^{q-1} b_{ij} g_j. \tag{2}$$

Now $h^i h^j = h^{i+j}$ if $i+j<q$ and $h^i h^j = h^{i+j-q} f$ if $i+j \geqq q$. Therefore

$$g_i g_j = g_{i+j} \prod_{k=1}^{s} x_k^{t_{k,i+j} - t_{k,i} - t_{k,j}} \text{ if } i+j<q$$

and

$$g_i g_j = g_{i+j-q} \Big(\prod_{k=1}^{s} x_k^{t_{k,i+j-q} - t_{k,i} - t_{k,j}} \Big) f = d g_{i+j-q} \Big(\prod_{k=1}^{s} x^{n_k + t_{k,i+j-q} - t_{k,i} - t_{k,j}} \Big), \text{ if } i+j \geqq q.$$

Therefore $a_i A_i = a_{i0} A_i + \sum_{j=1}^{q-1} a_{ij} A_i g_j$ and

$$b_i g_i = \sum_{p=0}^{i-1} d b_{i,q-i+p} \left(\prod_{k=1}^{s} x_k^{n_k + t_{k,p} - t_{k,i} - t_{k,q-i+p}} \right) g_p$$
$$+ \sum_{p=i}^{q-1} b_{i,p-i} \left(\prod_{k=1}^{s} x_k^{t_{k,p} - t_{k,i} - t_{k,p-i}} \right) g_p; \text{ [We take } t_{k,0} = 0 \text{ and } g_0 = 1].$$

Hence by (1) we get

$$A_0 = \sum_{i=1}^{H} a_{i0} A_i + d \sum_{i=1}^{q-1} b_{i,q-i} \left(\prod_{k=1}^{s} x_k^{n_k - t_{k,i} - t_{k,q-i}} \right) \tag{3}$$

Now $i n_k = t_{k,i} q + m_{k,i}$ and $(q-i) n_k = t_{k,q-i} q + m_{k,q-i}$ imply that

$$q n_k = (t_{k,i} + t_{k,q-i}) q + m_{k,i} + m_{k,q-i}$$

and hence $m_{k,i} + m_{k,q-i} = q$ and this in turn implies that $n_k - t_{k,i} - t_{k,q-i} = 1$. Therefore by (3) we get

$$A_0 = \sum_{i=1}^{H} a_{i0} A_i + d \left(\prod_{k=1}^{s} x_k \right) \sum_{i=1}^{q-1} b_{i,q-i}.$$

Comparing leading forms, we conclude that this is a contradiction since $s \geqq 2$.

Similarly for $i = 1, 2, \cdots, H$, there do not exist elements a_i, b_i in R for which $A_i = \sum_{i=0,1,\cdots,i-1,\, i+1,\cdots,H} a_i A_i + \sum_{i=1}^{q-1} b_i g_i$.

Now suppose, if possible, that there exist elements a_i, b_i in R^* such that $g_1 = h = \sum_{i=0}^{H} a_i A_i + \sum_{i=2}^{q-1} b_i g_i$. Using expressions (2) for a_i, b_i we get

$$1 = \sum_{i=0}^{H} a_{i1} A_i + d \sum_{i=2}^{q-1} b_{i,q-i+1} \left(\prod_{k=1}^{s} x_k^{n_k - t_{k,i} - t_{k,q-i+1}} \right). \tag{4}$$

Now $i n_k = t_{k,i} q + m_{k,i}$ and $(q - i + 1) n_k = t_{k,q-i+1} q + m_{k,q-i+1}$ and hence $q n_k + n_k = (t_{k,i} + t_{k,q-i+1}) q + m_{k,i} + m_{k,q-i+1}$. Therefore

$$q(n_k - t_{k,i} + t_{k,q-i+1}) = -n_k + m_{k,i} + m_{k,q-i+1}. \tag{5}$$

Replacing, if necessary, z by h_1, we may assume that $n_1 = 1$. Then $m_{1,i} = i$ and $m_{1,q-i+1} = q - i + 1$ so that $-n_1 + m_{1,i} + m_{1,q-i+1} = -1 + i + q - i + 1 = q$. Hence by (5), $n_1 - t_{1,i} - t_{1,q-i+1} = 1$. Hence by (4) we conclude that $1 \in M^*$ which is a contradiction.

Therefore the elements $A_0, A_1, \cdots, A_H$, $h = g_1$, of M^* are linearly independent modulo M^{*2} and hence

$$\dim(M^*/M^{*2}) > H + 1 = \dim R = \dim R^*.$$

Therefore R^* is non-regular.

THEOREM 7. *Let V be a projective normal model of an r-dimensional algebraic function field K/k of characteristic p. Let K^* be a cyclic extension of K of prime degree $q \neq p$, $(q > 1)$ and let V^* be a K^*-derived normal model of V and assume that k contains a primitive q-th root of unity. Let D be the branch locus on V of the rational transformation from V^* onto V. Let P be a simple point of V. Then the component of D passing through P is pure $(r-1)$-dimensional (or empty). If V is non-singular then D is pure $(r-1)$ dimensional (or empty).*

Proof. Let R be the quotient ring of P on V. Then R is a unique factorization domain. Let R^* be the integral closure of R in K^*. Then the discriminant ideal $D(R^*/R)$ of R^* over R defines the component of D passing through P. By Theorem 4, $D(R^*/R)$ is either the unit ideal or is a principal ideal other than the unit ideal. Hence the result.

7. Simultaneous resolution for quadratic and cubic cyclic extensions. The purpose of this section is to prove

THEOREM 8. *Let K be a two dimensional algebraic function field over a perfect ground field k of characteristic p and let K^* be a finite algebraic extension of K. Assume that either K^*/K is quadratic and $p \neq 2$ or that K^*/K is cubic cyclic, $p \neq 3$ and k contains a primitive cube root of unity. Then there exists a non-singular projective model V of K/k whose K^*-normalization is also non-singular and the branch locus on V is a non-singular curve (or is empty).*

Since k is perfect, K/k has non-singular models [4]; we remark that this the only thing for which the assumption of perfectness of k will be used and hence that the theorem remains true if the assumption of perfectness of k is replaced by the assumption that K/k has non-singular models.

Now let V be a non-singular projective model of K/k, and let V^* be a K^*-normalization of V. By Theorem 6, the branch locus on V is a curve D. Let $D_1, D_2, \cdots, D_n$ be the irreducible components of D. Let P be a singular point of D_1 and let V' be the surface obtained from V by applying a local quadratic transformation centered at P. Let D_i' be the curve on V' which corresponds to D_i and let L be the (non-singular fundamental) curve on V' which is the total image of P. Let V'^* be a K^*-normalization of V'. Then the branch locus on V' (for the transformation between V' and V'^*) consists either of $D_1', D_2', \cdots, D_n'$ or of $D_1', D_2', \cdots, D_n'$, L. In an obvious sense L intersects any of the curves D_i' normally.

Let $P_1, P_2, \cdots, P_N$ be the set of singularities of $D_1, D_2, \cdots, D_n$. Let V_1 be the surface obtained from V by applying a quadratic transformation centered at $(P_1, P_2, \cdots, P_N)$.[3] Let V_2 be the surface obtained from V_1 as V_1 is obtained from V. Let $V_1, V_2, V_3, \cdots$ be the sequence of surfaces defined in this manner. In view of Theorem, 4 of Part II of Zariski [11], it follows from the above remarks that for some t, all the irreducible components of the branch curve on V_t (for the transformation between V_t and its K^*-normalization) are non-singular. We may assume that this is so already for V, i.e. that the curves D_i are all non-singular.

Now let $Q_1, Q_2, \cdots, Q_M$ be the points of V which are common to more than one component of D, (i.e., the singular points of D). Let $V^{(1)}$ be the surface obtained from V by applying a quadratic transformation centered at $(Q_1, Q_2, \cdots, Q_M)$. Let $V^{(2)}$ be the surface obtained from $V^{(2)}$ as $V^{(1)}$ is obtained from V. Let $V^{(1)}, V^{(2)}, V^{(3)}, \cdots$ be the sequence of surfaces thus defined. Now suppose for instance that $Q_1 \in D_1$; let R be the quotient ring of Q_1 on V, let w be the rank two valuation of K/k having center at Q_1 composed with the real discrete valuation given by D_1 (w is unique since Q_1 is simple for D_1). Let R_i be the i-th quadratic transform of R along w. By Lemma 12 of [2], $\bigcup_{i=1}^{\infty} R_i = R_w$. In view of this, it is clear that we have proved the following proposition.

Proposition 4. *In the notation of Theorem 8, let V be a non-singular projective model of K/k, let $P_1, P_2, \cdots, P_N$ be the singular points of the branch curve D on V(of the transformation between V and its K^*-normalization). Let V_1 be the surface obtained from V by applying a quadratic transformation centered at $(P_1, P_2, \cdots, P_N)$. Let V_2 be the surface obtained from V_1 as V_1 is obtained from V. Let $V_1, V_2, V_3, \cdots$ be the sequence of surfaces thus defined. Let D_i be the branch curve on V_i for the transformation between V_i and its K^*-normalization. Then there exists an integer n such that for any $i \geqq n$, the only singularities of D_i are two-fold normal crossings)Definition 4 of [1]).*

Proposition 5. *Let (R, M) be an r-dimensional regular noetherian local domain, with quotient field K and let $(x_1, x_2, \cdots, x_r)$ be a basis of M. Let $K^* = K(z)$ with $z^q = dx_1^n$ where q is a prime number, n is a positive integer not divisible by q and d is a unit in R. Then there is only one local ring R^* in K^* lying above R and R^* is regular.*

---

[3] See part (2) of the proof of Theorem 3 of Section 4.

Proof. The proof follows from the argument given in the proof of Theorem 2 of [3].

In view of Proposition 4 the proof of Theorem 8 in the quadratic and the cubic cases follows respectively from Theorems 9 and Theorems 10 below.

THEOREM 9. *Let V be a non-singular surface with quotient field K/k of characteristic $\neq 2$ and K^* a quadratic extension of K. Let V^* be a K^*-normalization of V and assume that all the singularities $P_1, P_2, \cdots, P_n$ of the branch curve D on V for the transformation between V and V^* are ordinary double points. Let V' be the surface obtained from V by applying a quadratic transformation centered at $(P_1, P_2, \cdots, P_n)$ and let D' be the branch curve on V' for the transformation between V' and its K^*-normalization V'^*. Then V', V'^* and D' are all non-singular.*

Proof. Let P be one of the singular points of D and let (R, M) be the quotient ring of P on V. By Theorem 4 there exists a primitive element z of K^*/K such that $z^2 = dxy$ where (x, y) is a basis of M, d is a unit in R and $xy = 0$ is the local equation of D at P. We may replace x by dx and assume that $z^2 = xy$. Let (R_1, M_1) be an immediate two dimensional quadratic transform (Definition 3 of [2]) of R and let v be a valuation of K/k having center M_1 in R_1. We shall now show that there exists a primitive element z_1 of K^*/K such that either $z_1^2 = d_1$ or $z_1^2 = x_1$ where d_1 is a unit in R_1 and (x_1, y_1) is a basis of M_1 and in view of Theorem 4 and Proposition 5, the proof of the present theorem will then follow from this. *Case* 1, $v(x) > v(y)$. Let $y_1 = y$ and $x_1 = x/y$. Let $z = z/y$. Then (x_1, y_1) is a basis of M_1, z_1 is a primitive element of K^*/K and $z_1^2 = x_1$. *Case* 2, $v(x) < v(y)$. Because of symmetry in x and y the proof in this case follows from Case 1. *Case* 3, $v(x) = v(y)$. Let $d_1 = x/y$ and $z_1 = z/y$. Then d_1 is a unit in R_1, z_1 is a primitive element of K^*/K and $z_1^2 = d_1$.

THEOREM 10. *Let the notation be as in Theorem 9, except assume that K^*/K is cubic cyclic, k is of characteristic $\neq 3$ and contains a primitive cube root of unity. Let V'' be the surface obtained from V' as V' was obtained from V. Let D'' be the branch curve on V'' for the transformation between V'' and its K^*-normalization V''^*. Then V'', V''^* and D'' are all non-singular.*

Proof. Let P'' be an arbitrary point of D'', let P' be the point of V' which corresponds to P'' and let P be the point of V which corresponds to P'. Then $P' \in D'$ and $P \in D$. Let (R_2, M_2) and (R_1, M_1) be the quotient rings of P'' and P' on V' and V'' respective and let v be a valuation of K/k

having center M_2 in R_2. In view of Theorem 4 and Proposition 5, it is enough to show that P'' is a simple point of D''.

If P' is a simple point of D' then $R_2 = R_1$ and P'' is a simple point of D''. Now assume that P' is singular for D'. Then P' and P are ordinary double points of D' and D respectively. By Theorem 4, we can find a primitive element z of K^*/K such that either (Case 1) $z^3 = dxy^2$ or (Case 2) $z^3 = dx^2y$ or (Case 3) $z^3 = dxy$ where d is a unit in R and (x, y) is a basis of M.

Suppose, if possible, that we have Case 1, i.e., $z^3 = dxy^2$. We have the following three subcases. *Subcase* 1.1: $v(x) < v(y)$. Let $z_1 = z/x$, $x_1 = x$, and $y_1 = y/x$. Then z_1 is a primitive element of K^*/K, (x_1, y_1) is a basis of M_1 and $z_1{}^3 = dy_1{}^2$ and hence by Theorem 4, P' is a simple point of D' which is a contradiction. *Subcase* 1.2: $v(x) > v(y)$. Let $z_1 = z/y$, $x_1 = x/y$ and $y_1 = y$. Then z_1 is a primitive element of K^*/K, (x_1, y_1) is a basis of M_1 and $z_1{}^3 = dx_1$ and hence by Theorem 4, P' is a simple point of D' which is a contradiction. *Subcase* 1.3: $v(x) = x(y)$. Then $z_1 = z/y$ is a primitive element of K^*/K and $z_1{}^3 = e$ where $e = dx/y$ is a unit in R_1; hence P' is not a point of D' which is a contradiction.

Thus Case 1 is not possible and similarly Case 2 is not possible either. Thus we are left with Case 3, i.e., we have $z = dxy$. Suppose, if possible, that $v(x) = v(y)$. Let $x_1 = x$ and $d_1 = dy/x$. Then x_1 is part of a basis (x_1, y_1) of M_1, d_1 is a unit in R_1 and $z^3 = d_1x_1{}^2$; hence by Theorem 4, P' is a simple point of D' which is a contradiction. Therefore $v(x) \neq v(y)$. Because of symmetry in x and y, it is enough to complete the argument in case $v(x) > v(y)$. Let $y_1 = y$ and $x_1 = x/y$. Then (x_1, y_1) is a basis of M_1 and $z^3 = dx_1y_1{}^2$. The argument which implied the impossibility of Case 1 now tells us that $z^3 = dx_1y_1{}^2$ implies that P'' is a simple point of D''.

8. Simultaneous nonresolvability for cyclic extensions of minimal models.

DEFINITION. *Let K/k be an algebraic function field and let K^* be a finite algebraic extension of K. If there does not exist any non-singular projective model of K/k with a non-singular K^*-normalization then we shall say that (K, K^*) is simultaneously non-resolvable.*

Now let (R, M) be a regular two dimensional (noetherian) local domain with quotient field K. Let (x_0, y_0) be a basis of M. Let a be a fixed positive integer and let $b = a + 1$. Let

$$
(1)\qquad
\begin{array}{ll}
x_1 = x_0/y_0, & y_1 = y_0; \\
x_2 = x_1/y_1, & y_2 = y_1; \\
\cdots & \\
x_{b-1} = x_{b-2}/y_{b-2}, & y_{b-1} = y_{b-2}; \\
x_b = x_{b-1}, & y_b = y_{b-1}/x_{b-1}; \\
x_{b+1} = x_b/y_b, & y_{b+1} = y_b; \\
x_{b+2} = x_{b+1}/y_{b+1}, & y_{b+2} = y_{b+1}; \\
\cdots & \\
x_{2b-1} = x_{2b-2}/y_{2b-2}, & y_{2b-1} = y_{2b-2}; \\
x_{2b} = x_{2b-1}, & y_{2b} = y_{2b-1}/x_{2b-1}. \\
\cdots &
\end{array}
$$

i. e. for $n = 0, 1, 2, \cdots$ we have set

$$
(2)\qquad
\begin{array}{lll}
x_{nb+i} = x_{nb+i-1}/y_{nb+i-1}, & y_{nb+i} = y_{nb+i-1} & \text{for } i = 1, 2, \cdots, b-1 \\
\text{and } x_{nb+b} = x_{nb+b-1}, & y_{nb+b} = y_{nb+b-1}/x_{nb+b-1}; &
\end{array}
$$

in other words

$$
(3)\qquad
\begin{array}{lll}
x_{nb+i} = x_{nb}/y_{nb}{}^{i}, & y_{nb+i} = y_{nb} & \text{for } i = 1, 2, \cdots, b-1, \\
\text{and } x_{(n+1)b} = x_{nb}/y_{nb}{}^{b-1}, & y_{(n+1)b} = y_{nb}{}^{b}/x_{nb}. &
\end{array}
$$

Then there is a unique two dimensional h-th quadratic transform (Definition 3 of [2]) (R_h, M_h) of (R, M) such that $M_h = (x_h, y_h)R_h$. More explicitly we can define (R_h, M_h) either by induction as:

For $n = 0, 1, 2 \cdots$ let

$$S_{nb+i} = R_{nb+i-1}[x_{nb+i}], \quad N_{nb+i} = (x_{nb+i}, y_{nb+i})S_{nb+i} \text{ for } i = 1, 2, \cdots, b-1,$$

and

$$S_{nb+b} = R_{nb+b-1}[y_{nb+b}], \qquad N_{nb+b} = (x_{nb+b}, y_{nb+b})S_{nb+b};$$

and then

$$R_{nb+i} = (S_{nb+i})_{N_{nb+i}}, \quad M_{nb+i} = N_{nb+i}R_{nb+i} \text{ for } i = 1, 2, \cdots, b.$$

Or directly as:

$$S_t^* = R[x_t, y_t], \quad N_t^* = (x_t, y_t)S_t^*, \quad R_t = (S_t^*)_{N_t^*} \text{ and } M_t = N_t^* R_t^*.$$

Anyway, R_{t+1} is an immediate quadratic transform of R_t for $t = 0, 1, 2, \cdots$ where $R_0 = R$.

By Lemma 12 of [2], $\bigcup_{i=1}^{\infty} R_i = R_v$ where v is a valuation of K having center M_i in R_i. It follows from the definition of the sequence (x_1, y_1),

$(x_2, y_2), \cdots$ that $v(x)$ and $v(y)$ are comparable [i.e., there exist positive integers A and B such that $Av(x) > v(y)$ and $Bv(y) > v(x)$] and that (the real number) $Q = v(x)/v(y)$ is an irrational number given by the periodic continued fraction expansion (of period two):

$$Q = a + 1/(1 + 1/(a + 1/(1 + 1/(a + \cdots.$$

Hence, in view of the results of [2], v is a real valuation of rational rank two and $R_v/M_v = R/M = R_i/M_i$ for any i. If we set $v(y) = 1$ and $v(x) = Q$, then the value group of v consists of the real numbers of the form $m + nQ$ where m and m are arbitrary integers.

THEOREM 11. *Let (R, M) be a two dimensional regular (noetherian) local domain with quotient field K of characteristic p and let (x, y) be a basis of M. Let q be a prime number such that $q > 3$ and $q \neq p$. Let $a = q - 4$, $b = a + 1 = q - 3$, $x_0 = x$ and $y_0 = y$. Let v be the valuation of K as constructed above. Let K^* be the root field over K of the polynomial $Z^q - xy^2$. Let (S, N) be any two dimensional regular (noetherian) local domain with quotient field K such that $R_v \supset S \supset R$, $M_v \cap S = N$ and $M = R \cap N$. Then there is only one local ring S^* in K^* lying above S and S^* is not regular.*

Proof. We shall use the notation introduced above in the construction of v. By Theorem 3 of [2] there exists an integer h such that $S = R_h$ and $N = M_h$. Thus we have to prove that for any integer t, there is only one local ring R_t^* in K^* lying above and R_t^* is non-regular.

Let H be a root of $Z^q - xy^2 = 0$. Then H is a primitive element of K^*/K and K^*/K is cyclic of degree q. Let $F = xy^2$. Let us rewrite equations (3) as:

For $n = 0, 1, 2, \cdots$ and $i = 1, 2, \cdots, b - 1$,

$$(4) \qquad \begin{aligned} &x_{nb} = (x_{nb+i})(y_{nb+i})^i = (x_{nb+b})^b (y_{nb+b})^{b-1}, \qquad y_{nb} = y_{nb+i} = x_{nb+b} y_{nb+b}, \\ &x_{nb+b-1} = x_{nb+b} \qquad y_{nb+b-1} = x_{nb+b} y_{nb+b}. \end{aligned}$$

Let $F_i = F$ and $H_i = H$ for $i = 0, 1, \cdots, b$. Then for $i = 1, 2, \cdots, b - 1$ we have $F_i = xy^2 = (x_i y_i^i)(y_i)^2 = x_i y_i^{i+2}$ with $0 < i + 2 \leqq (b - 1) + 2 = b + 1 < q$, and $F_b = (x_b^b y_b^{b-1})(x_b y_b)^2 = x_b^{b+2} y_b^{b+1}$ with $0 < b + 2 < q$ and $0 < b + 1 < q$.

For $i = 1, 2, \cdots, b - 1$ let $F_{b+i} = F_b (y_{b+i})^{-iq}$ and $H_{b+i} = H_b (y_{b+i})^{-i}$. Then H_{b+i} is a primitive element of K^*/K and

$$\begin{aligned}(H_{b+i})^q &= F_{b+i} = F_b(y_{b+i})^{-iq} = x_b^{b+2}y_b^{b+1}(y_{b+i})^{-iq}\\ &= (x_{b+i}y_{b+i}{}^i)^{b+2}(y_{b+i})^{b+1}(y_{b+i})^{-iq}\\ &= (x_{b+i})^{b+2}(y_{b+i})^{i(b+2)+b+1-i(b+3)} = (x_{b+i})^{b+2}(y_{b+i})^{b+1-i}\end{aligned}$$

and $0 < b+2 < q$, $0 < b+1-i < q$.

In particular, substituting $i = \mathrm{b}-1$ we get

$$(H_{2b-1})^q = F_{2b-1} = (x_{2b-1})^{b+2}(y_{2b-1})^2.$$

Let $F_{2b} = F_{2b-1}(x_{2b})^{-q}$ and $H_{2b} = H_{2b-1}(x_{2b})^{-1}$. Then H_{2b} is a primitive element of K^*/K and

$$\begin{aligned}(H_{2b})^q = F_{2b} = F_{2b-1}(x_{2b})^{-q} &= (x_{2b-1})^{b+2}(y_{2b-1})^2(x_{2b})^{-q}\\ &= (x_{2b})^{b+2}(x_{2b}y_{2b})^2(x_{2b})^{-q} = (x_{2b})^{b+4-q}(y_{2b})^2 = x_{2b}(y_{2b})^2\end{aligned}$$

i. e.,

$$(H_{2b})^q = F_{2b} = x_{2b}(y_{2b})^2. \tag{5}$$

Let $F_{2b+i} = F_{2b}$, $H_{2b+i} = H_{2b}$ for $i = 1, 2, \cdots, b$; $F_{3b+i} = F_{3b}(y_{3b+i})^{-iq}$, $H_{3b+i} = H_{3b}(y_{3b+i})^{-i}$ for $i = 1, 2, \cdots, b-1$; and $F_{4b} = F_{4b-1}(x_{4b})^{-q}$, $H_{4b} = H_{4b-1}(x_{4b})^{-1}$. Then in view of (5), the above considerations at once tell us that H_{2b+i} is a primitive element of K^*/K for $i = 1, 2, \cdots, 2b$; and

$$\begin{aligned}&(H_{2b+i})^q = F_{2b+i} = x_{2b+i}(y_{2b+i})^{i+2} \text{ for } i = 1, 2, \cdots, b-1;\\ &(H_{3b+i})^q = F_{3b+i} = (x_{3b+i})^{b+2}(y_{3b+i})^{b+1-i} \text{ for } i = 0, 1, \cdots, b-1;\\ &(H_{4b})^q = F_{4b} = x_{4b}(y_{4b})^2.\end{aligned}$$

And so on $\cdots$, i. e., by induction we define for $n = 0, 1, 2, \cdots$

$$F_{2nb+i} = F_{2nb}, \quad H_{2nb+i} = H_{2nb} \text{ for } i = 1, 2, \cdots, b;$$
$$F_{2nb+b+i} = F_{2nb+b}(y_{2nb+b+i})^{-iq}, \quad H_{2nb+b+i} = H_{2nb+b}(y_{2nb+b+i})^{-i} \quad \text{for } i = 1, 2, \cdots, b-1;$$

and

$F_{2nb+2b} = F_{2nb+2b-1}(x_{2nb+2b})^{-q}$, $H_{2bn+2b} = H_{2nb+2b-1}(x_{2nb+2b})^{-1}$; and we conclude that for any non-negative t, H_t is a primitive element of K^*/K and

$H_t{}^q = F_t = x_t{}^{A_t}y_t{}^{B_t}$ with $0 < A_t < q$ and $0 < B_t < q$. [Observe that $A_{2nb+i} = A_i$, $B_{2nb+i} = B_i$ for $n = 0, 1, 2, \cdots$ and $i = 1, 2, \cdots, 2b$]. Hence by Theorems 4 and 6 it follows that there is only one local ring R_t^* in lying above R_t and R_t^* is not regular. Q. E. D.

Now let K/k be a two dimensional algebraic function field and recall that a non-singular projective model V of K/k is called a *minimal model* of K/k if W is a non-singular projective model of K/k implies that the

(uniquely determined) birational transformation from W onto V is everywhere regular on W. In [14] Zariski has proved (at least when k is algebracically closed) that K/k has a minimal model if and only if there does not exist any intermediate field between k and K of which K is a simple transcendental extension.

THEOREM 12. *Let K/k be a two dimensional algebraic function field of characteristic p and assume that K/k has a minimal model. Let q be a prime number such that $q > 3$ and $q \neq p$. Assume that k contains a primitive q-th root of unity. Then there exist q-cyclic extensions (lots of them) K^* of K such that (K, K^*) is simultaneously non-resolvable.*

Proof. Let V be the minimal model of K/k. Let (R, M) be the quotient ring of a point of V. Let (x, y) be a basis of M. Let K^* be a root field over K of the polynomial $Z^q - xy^2$. It then follows by Theorem 11 that (K, K^*) is simultaneously non-resolvable.

9. Partial generalizations to higher varieties. In this section we plan to give partial generalizations to higher varieties of the theorems proved in Sections 2, 3 and 7 for surfaces. We start off by deducing the following consequence (which is the weak local uniformization theorem for real valuations for an algebraically closed ground field of characteristic zero) of a result proved by Zariski in his uniformization paper [9].

THEOREM 13. *Let P be a simple point on an r-dimensional algebraic variety V with function field K/k where the ground field k is algebraically closed of characteristic zero. Let (R, M) be the quotient ring of P on V. Let v be a zero dimensional real valuation of K/k having center P on V and let u be a non-zero element of R. Then there exists an anti-regular transform V^* of V obtained from V by a sequence of monoidal transformations such that there is a minimal basis $(x_1^*, x_2^*, \cdots, x_r^*)$ of the maximal ideal M^* of the quotient ring R^* of the center P^* of v on V^* for which $u = x_1^{*a_1} x_2^{*a_2} \cdots x_h^{*a_h} d$ where $a_1, a_2, \cdots, a_h$ are non-negative integers, d is a unit in R^* and $v(x_1), v(x_2), \cdots, v(x_h)$ are rationally independent.*

It is clear that this theorem will follow from the following slightly more general proposition.

PROPOSITION 6. *Let (R, M) be an r-dimensional $(r > 1)$ regular (noetherian) local domain with quotient field K. Assume that R contains a subfield k which is mapped isomorphically onto the residue field R/M by the canonical homomorphism and assume that k is algebraically closed of*

characteristic zero. Let $(x_1, x_2, \cdots, x_r)$ *be a basis of* M. *Let* v *be a real valuation of* K/k *having center* M *in* R *and of* R*-dimension zero,* (i. e., $k = R_v/M_v$). *Let* u *be a non-zero element in* M. *Then there exists a regular* r*-dimensional local domain* (R^*, M^*) *obtained from* R *by a sequence of monoidal transformations with respect to the basis* $(x_1, x_2, \cdots, x_r)$ *and the field* k *(see definition below) and a basis* $(x_1^*, x_2^*, \cdots, x_r^*)$ *of* M^* *such that* $u = dx_1^{*a_1}x_2^{*a_2}\cdots x_h^{*a_h}$ *where* d *is a unit in* R^*; $a_1, a_2, \cdots, a_h$ *are positive integers and* $v(x_1), v(x_2), \cdots, v(x_h)$ *are rationally independent.*

Proof. For $i = 1, 2, \cdots, r$ let $f_i(X_1, X_2, \cdots, X_r)$ be elements of $k[X_1, X_2, \cdots, X_r]$ of leading degree one such that if $g_{i1}X_1 + g_{i2}X_2 + \cdots + g_{ir}X_r$ with g_{ij} in k is the leading form of f_i then the r by r determinant $(g_{ij}) \neq 0$. Let $y_i = f_i(x_1, x_2, \cdots, x_r)$. Then $(y_1, y_2, \cdots, y_r)$ is a basis of M. Assume that $v(y_1) < v(y_i)$ for $i = 2, 3, \cdots, t$ with $t \leqq r$. Let $z_1 = y_1$, $z_i = y_i/y_1$ for $i = 2, 3, \cdots, t$ and $z_i = y_i$ for $i = t+1, t+2, \cdots, r$. Let $S = R[z_1, z_2, \cdots, z_r]$, $N = S \cap M_v$ and $R_1 = S_N$ and $M_1 = NR_1$. Then (R_1, M_1) is a regular r-dimensional local domain, $(z_1, z_2, \cdots, z_r)$ is a basis of M_1, $R_v \supset R_1 \supset R$, $M_v \cap R_1 = M_1$ and $M_1 \cap R = M$. We shall say that (R_1, M_1) is a first (or an immediate) monoidal transform of R along v with respect to the basis $(x_1, x_2, \cdots, x_r)$ and the field k. Let (R_2, M_2) be an immediate monoidal transform of R_1 along v with respect to the basis $(z_1, z_2, \cdots, z_r)$ and the field k. Then we shall say that R_2 is a second monoidal transform of R along v with respect to the basis $(x_1, x_2, \cdots, x_r)$ and the field k and so on.

Now let $A = k[x_1, x_2, \cdots, x_r]$, $Q = (x_1, x_2, \cdots, x_r)A$, $R' = A_Q$, $M' = QR'$, and let v' be the restriction of v to $k(x_1, x_2, \cdots, x_r)$. It is a direct consequence of what Zariski has proved in parts B, CI and CII of [9] that the present proposition holds for (R', M'); this special case of the proposition will be referred to as Z.

Let $E = \min(v(x_1), v(x_2), \cdots, v(x_r))$ and let $v(u) = D$. Fix a positive integer t such that $tE > D$. We can write $u = u_1 + u'$ where u_1 is a polynomial in $k[x_1, x_2, \cdots, x_r]$ of degree less than t and $u' \in M^t$. Then $v(u_1) = v(u) = D$ and $u = u_1 + b_2u_2 + \cdots + b_Tu_T$ where $b_i \in R$ and $u_2, u_3, \cdots, u_T$ are monomials of degree t in $x_1, x_2, \cdots, x_r$. Applying Z to the product $u_1u_2\cdots u_T$ we conclude that there exists an n-th monoidal transform (R'^*, M'^*) of R' along v' with respect to the basis $(x_1, x_2, \cdots, x_r)$ and the field k such that $u_i = d_ix_1^{*a_{i1}}x_2^{*a_{i2}}\cdots x_H^{a_{iH}}$ for $i = 1, 2, \cdots, T$ where the a_{ij} are non-negative integers, d_i is a unit in R'^*, $(x_1^*, x_2^*, \cdots, x_r^*)$ is a basis of M'^* and $v'(x_1^*), v'(x_2^*), \cdots, v'(x_H^*)$ are rationally independent.

Since $v'(u_i) > v'(u_1)$ for $i = 2, 3, \cdots, T$; invoking Theorem 2 of [9] we may arrange matters so that $a_{1j} < a_{ij}$ for $j = 1, 2, \cdots, H$ and $i = 2, 3, \cdots, T$. Let $S = R[x_1^*, x_2^*, \cdots, x_r^*]$, $N = (x_1^*, x_2^*, \cdots, x_r^*)$, $R^* = S_N$ and $M^* = NR^*$. Then evidently (R^*, M^*) is an n-th monoidal transform of R along v with respect to the basis $(x_1, x_2, \cdots, x_r)$ and the field k and $(x_1^*, x_2^*, \cdots, x_r^*)$ is a basis of M^*. Also it is clear that

$$u = dx_1^{*a_{11}} x_2^{*a_{12}} \cdots x_H^{*a_{1H}}$$

where d is a unit in R^*. Rearrange the $x_1^*, x_2^*, \cdots, x_H^*$ so that $a_{1i} \neq 0$ for $i = 1, 2, \cdots, h$ and $a_{1i} = 0$ for $i > h$. Setting $a_{1i} = a_i$ for $i = 1, 2, \cdots, h$ we obtain the required result.

PROPOSITION 7. *Let the notation be as in Proposition 6 and let w be an element of R_v. Then there exists R^* as described in Proposition 6 such that $w \in R^*$.*

Proof. Let $w = w_1/w_2$ with $w_1, w_2 \in R^*$. Applying Proposition 6 to the product $w_1 w_2$ and invoking Theorem 2 of Zariski [9] we obtain the result.

Now we give a partial generalization of Theorem 1 (Section 2) to higher varieties.

THEOREM 1A. *Let K be an n-dimensional algebraic function field over an algebraically closed ground field k of characteristic zero. Let K^* be a finite algebraic extension of K and let v^* be a zero dimensional real valuation of K^*/k. Let r be the rational rank of v^*. Then there exists a regular local ring (R^*, M^*) which is the quotient ring of the center of v^* on a projective model V^* of K^*/k such that if we set $S = R^* \cap K$, $Q = (M^* \cap K)$ and $Q^* = QR^*$ then we have the following: (1) S is normal domain with quotient field K and Q is the unique maximal ideal in S; (2) if $r = 1$ then rank $Q^* \geqq 2$; (3) if $r > 1$ then rank $Q^* \geqq r$; (4) if $r = n$ then V^* is a K^*-normalization of a projective normal model of K/k.*

Proof. Since v can be uniformized [9], there exists a regular local ring (R^*, M^*) which is the quotient ring of the center of v^* on a projective normal model of K^*/k. To arrange matters so that the transcendence degree of S over k is n we may either use the argument used in Theorem 1 together with Proposition 7 or we may invoke the fact that Zariski has proved the local uniformization theorem for v in projective form [9]. The rest of (1) is obvious. In view of Theorem 13, (2) follows by the argument used in Case 1 in the proof of Theorem 1.

Now assume that $r > 1$. Then r is also the rational rank of the K-

restriction of V^*. Fix elements $z_1, z_2, \cdots, z_r$ in K such that $v^*(z_1), v^*(z_2), \cdots, v^*(z_r)$ are rationally independent and positive. Either by Proposition 7 or by the projective form of the uniformization theorem [9], we may assume that the elements z_i are in R^*. Applying Theorem 13 to the product $z_1 z_2 \cdots z_r$ we may assume that $z_i = d_i x_1^{a_{i1}} x_2^{a_{i2}} \cdots x_h^{a_{ih}}$ where $(x_1, x_2, \cdots, x_n)$ is a basis of M^*, d_i is a unit in R^*, a_{ij} is a non-negative integer, and $v^*(x_1), v^*(x_2), \cdots, v^*(x_h)$ are rationally independent. Since $v^*(z_1), v^*(z_2), \cdots, v^*(z_r)$ are rationally independent we must have $h = r$ and for the r by r determinants $D = (a_{ij})$ we have $D \neq 0$. Hence there exist integers n_{ij} and units e_j in R^* such that $e_j x_j^D = z_1^{n_{1j}} z_2^{n_{2j}} \cdots z_r^{n_{rj}}$ for $j = 1, 2, \cdots, r$. Therefore the elements $x_1, x_2, \cdots, x_r$ belong to the radical of Q^* and hence rank $Q^* \geqq r$. This proves (3) and (4) follows from (3) and Proposition 1 of [4].

Proposition 3 of Section 3 can be generalized as follows:

PROPOSITION 3A. *Let K be an r-dimensional algebraic function field of characteristic zero and K^* a galois extension of K. Let v^* be a zero dimensional real valuation of K^* and let v be the K-restriction of v^*. Then there exists a regular local ring R which is the quotient ring of the center of v on a projective normal model V of K such that if R^* denotes the quotient ring of the center of v^* on a K^*-normalization of V then the splitting field of R^* over R coincides with the splitting field of v^* over v.*

Proof. The argument used in the proof of Theorem 1 of [4] tells us that in order that what is required should happen, it is sufficient that a certain finite number of elements of R_v be contained in R. Now this can be arranged either by using Proposition 7 or simply by invoking the fact that Zariski has proved the local uniformization theorem for v in projective form [9].

Now we generalize Theorem 2 (local simultaneous resolution for rational valuations) to higher varieties.

THEOREM 2A. *This is the same as Theorem 2 except let the dimension of K/k be arbitrary.*

Proof. In view of Proposition 3A and Theorem 13 the proof is the same as that of Theorem 2.

The results of Sections 7 and 8 (simultaneous resolvability or non-resolvability) and the partial generalizations of the results of Section 7 to higher varieties seems to be somewhat related to (and hence they might throw some light on) the possible approach to the problem of resolution of singu-

larities for higher varieties (over an algebraically closed ground field of characteristic zero) which has been proposed by Zariski in [13]. Hence before giving these generalizations of the results of Section 7, we wish to make a remark concerning this proposed method of Zariski. This method calls for proving the following stronger conjectural form of the resolution problem.

THEOREM Z. (Theoréme de Réduction of pages 4-5 of [13]). *If V_r is a pure r-dimensional variety regularly immersed in the direct product F_{r+1} of $r+1$ projective lines (over the field of complex numbers), then there exists an anti-regular transformation f of F_{r+1} such that $f(F_{r+1})$ is non-singular and such that all the singularities of the total transform $f[V_r]$ of V_r are 2-fold normal crossings (for the definition of an s-fold normal crossings see* [1]; *what we call a "2-fold normal crossing" is in Zariski's terminology simply a "normal crossing").*

We shall prove below that for any r there exists a V_r satisfying the conditions of Theorem Z such that if f is an anti-regular transformation of F_{r+1} with the properties: (1) $f(F_{r+1})$ is non-singular and (2) all the singularities of $f[V_r]$ are normal crossings; then $f[V_r]$ has a singularity which is an $(r+1)$-fold normal crossing.

Let $t=r+1$ and let $x_1, x_2, \cdots, x_t$ be t algebraically independent elements over an algebraically closed ground field k of characteristic zero. Let $y_i = x_i^p$ where p is a prime number. Let $K = k(y_1, y_2, \cdots, y_t)$ and let $K^* = k(x_1, x_2, \cdots, x_t)$. Let $q_1, q_2, \cdots, q_t$ be rationally independent positive real numbers. Now there is a unique zero dimensional valuation v^* of K^*/k such that $v^*(x_i) = q_i$ for $i = 1, 2, \cdots, t$. Let v be the K-restriction of v^*. Let G be the galois group of K^*/K. Then G is the direct product of t cyclic groups of order p and *hence G is not a direct product of s cyclic groups for any $s < t$.* It is clear that the ramification index of v^* over v is (pt) and hence G is also the splitting group of v^* over v.

In a natural way we can consider K/k to be the function field of the direct product F_t of t projective lines over k. Let F^* be a K^*-normalization of F_t and consider the branch locus on F_t of the transformation between F_t and F^*. By a Theorem of Zariski (Theorem 1 of [1]) this branch locus is a pure r-dimensional variety and we let it be our V_r. The conditions of immersion are readily adjusted (for instance: roughly speaking, take a line not on $y_1 y_2 \cdots y_t = 0$ as a factor line of F_t etc.). Let f be an anti-regular transformation of F_t satisfying conditions (1) and (2) above. Let $F' = f(F_t)$ and F'^* be a K^*-normalization of F'. Let D be the branch locus on F' of

the transformation between F' and F'^{*} and let $V' = f[V_r]$. Then $D \subset V'$ and hence all the singularities of D are normal crossings. Let P be the center of v on F' and let P^* be the center of v^* on F'^*. Then $G \supset G(P^*:P) \supset$ (the splitting group of v^* over v), see [3]. Therefore $G(P^*:P) = G$. Therefore P is on D and hence P is an s-fold normal crossing of D for some s. By Theorem 2 of [1], $s = t$. Therefore P is a t-fold normal crossing also of V'. Q. E. D.

This shows that Theorem Z is not true in its present form. Also we have that, Theorem Z must be weakened at least to the point as to read: . . . every singularity of $f[v_r]$ is an s-fold normal crossing with $s \leqq r+1$ (s depending on the singularity).

Now we proceed to generalize Theorems 9 and 10 of Section 7 to higher varieties and we shall closely follow the proofs of these theorems.

THEOREM 9A. *Let K be an r-dimensional algebraic function field of characteristic $\neq 2$ and K^* a quadratic extension of K. Let V be a non-singular projective model of K and let V^* be K^*-normalization of V. Let D be the branch locus on V for the transformation between V and V^*. By Theorem 7 we have that D is pure $r-1$ dimensional. Assume that the only singularities of D are s-fold normal crossings with $s \leqq r$ (s depends on the singularity). Then there exists a non-singular model V_1 of K obtained from V by a sequence of monoidal transformations such that the K^*-normalization V_1^* of V_1 and the branch locus D_1 on V_1 are both non-singular.*

Proof. (We only give the additional considerations to supplement the proof of Theorem 9.) If $P \in D$ is an s-fold normal crossing of D then we shall denote the integer s by $s(P)$. Let $t = \max_{P \in D}[s(P)]$. If $t = 1$ there is nothing to prove, so assume that $t > 1$. Fix $P \in D$ with $s(P) = t$. Let (R, M) be the quotient ring of P on V and let $(x_1, x_2, \cdots, x_r)$ be a basis of M such that $x_1 x_2 \cdots x_t = 0$ is a local equation of D at P. There exists a primitive element z of K^*/K such that $z^2 = d x_1 x_2 \cdots x_t$ where d is a unit in R. Let v be a zero dimensional valuation of K/k having center P on V and let (R', M') be the quotient ring of the center P' of v on the variety V' obtained from V by a monoidal transformation centered at the $(r-2)$-dimensional subvariety W with local P-equation $x_{t-1} = 0 = x_t$. Let $y_i = x_i$ for $i = 1, 2, \cdots, t-2$. If $v(x_{t-1}) \neq v(x_t)$, let $y_{t-1} = x_t/x_{t-1}$ and $z_1 = z/x_{t-1}$ in case $v(x_{t-1}) < v(x_t)$, and let $y_{t-1} = x_{t-1}/x_t$ and $z_1 = z/x_t$ in case $v(x_t) < v(x_{t-1})$; then z_1 is a primitive element K^*/K with $z_1^2 = d y_1 y_2 \cdots y_{t-1}$ and $(y_1, y_2, \cdots, y_{t-1})$ is part of a minimal basis of M'. If $v(x_{t-1}) = v(x_t)$, then letting $z_1 = z/x_t$ we have that $z_1^2 = d_1 y_1 y_2 \cdots y_{t-2}$, where d_1 is a unit

in R' and $(y_1, y_2, \cdots, y_{t-2})$ is part of a minimal basis of M'. Thus in either case $s(P') < t$. Similar things happen to any other point on the subvariety W. Hence if there is still a point Q on V' with $s(Q) = t$ then Q does not lie on the total transform of W on V'. Now we treat Q on V' as we treated P on V and so on. Eventually we achieve a reduction in t. We continue till t reduces to 1.

THEOREM 10A. *Let the assumption be as in Theorem* 9A *except assume that K is of characteristic $\neq 3$, K^* is a cubic cyclic extension of K and that K contains a primitive cube root of unity. Then we have the same conclusion as in Theorem* 9A.

Proof. Combine the arguments used in the proofs of Theorems 10 and 9A.

COLUMBIA UNIVERSITY.

REFERENCES

[1] S. Abhyankar, "On the ramification of algebraic functions," *American Journal of Mathematics*, vol. 77 (1955), pp. 575-592.

[2] ———, "On the valuations centered in a local domain," *American Journal of Mathematics*, vol. 78 (1956), pp. 321-348.

[3] ———, "Local uniformization on algebraic surfaces over ground fields of charac- $p \neq 0$," *Annals of Mathematics*, vol. 63 (1956), pp. 491-526.

[4] ———, "On the field of definition of a nonsingular birational transform of an algebraic surface," *Annals of Mathematics* (to appear).

[5] Wolfgang Krull, "Galoissche Theorie bewerteter Körper," *Sitzungsberichte der Bayerischen Akademie der Wissenschaften*, München (1930), pp. 225-238.

[6] T. Matsusaka, "A note on my paper, 'Some theorems on abelian varieties'," *National Science Report of the Ochanomizu University*, vol. 5 (1954), pp. 21-23.

[7] Van der Waerden, *Modern Algebra*, vol. II, New York, 1950.

[8] O. Zariski, "The reduction of the singularities of an algebrais surface," *Annals of Mathematics*, vol. 40 (1939), pp. 639-689.

[9] ———, "Local uniformization on algebraic varieties," *ibid.*, vol. 41 (1941), pp. 852-896.

[10] ———, "Foundations of a general theory of birational correspondences," *Transactions of the American Mathematical Society*, vol. 53 (1943), pp. 490-542.

[11] ———, "Reduction of the singularities of algebraic three-dimensional varieties," *Annals of Mathematics*, vol. 45 (1944), pp. 472-542.

[12] ———, "Theory and applications of holomorphic function on algebraic varieties over arbitrary ground fields," *Memoirs of the American Mathematical Society*, No. 5, 1951.

[13] ———, "Le problème de la réduction des singularités d'une variété algébrique," *Bulletin des Sciences Mathématiques*, vol. 78 (1954), pp. 1-10.

[14] ———, "On minimal models of algebraic surfaces," forthcoming.

SOME ALGEBRAIC NUMBER THEORY ESTIMATES BASED ON THE DEDEKIND ETA-FUNCTION.*

By Harvey Cohn.

1. Introduction. Recent papers in this Journal [1], [2] attesc to the continued interést in bounds on class number, regulator, etc., of algebraic fields. One special type, known as the *pure cubic* field (see Section 2 below) has the advantage that the zeta-function residue reduces to a finite expression involving the Dedekind eta-function. From this expression, by using the modular sub-diivsion, we can obtain the estimate (see Section 3 below).

$$h \log \epsilon = O(|d|^{\frac{1}{2}} \log |d| \log\log |d|), \tag{1.1}$$

where d is the field discriminant, h is the class number, and $\epsilon(>1)$ is the fundamental unit. The estimate (1.1) is a slight improvement over Landau's estimate [8], $O(|d|^{\frac{1}{2}}\log^2|d|)$ under broader choice of field. One can obtain the further estimate (see Section 4 below).

$$h = O(|d|^{\frac{1}{2}} \log\log |d|) \tag{1.2}$$

which is a somewhat greater improvement over Landau's [7] $O(|d|^{\frac{1}{2}}\log^2|d|)$. A more direct application of the technique would be bounds on certain modular invariants which are included for purposes of illustration. The main significance of this paper, however, is an indication of the direct usefulness of modular functions for cubic field estimates.

2. Notation and formulas. To discuss the pure cubic field $K_{a,b}$, let

a, b	be two relatively prime square-free positive integers, $(ab \neq 1)$,
$K_{a,b}$	be the field generated by the real root $(ab^2)^{1/3}$,
h	be the class number of $K_{a,b}$,
$\epsilon(>1)$	be the fundamental unit of $K_{a,b}$,
k	be ab (or $3ab$) according as $a^2 - b^2 \equiv 0$ (or $\not\equiv 0$) mod 9,
d	be $-3k^2$, the discriminant of $K_{a,b}$.

Then according to the famous formula of Dedekind ([3]; p. 228)

$$\epsilon^h = \Pi \mathrm{H}(\omega_1)/\Pi \mathrm{H}(\omega_0), \tag{2.11}$$

* Received March 21, 1956.

From the shape of the fundamental domain (2.42),

$$(2.44)\qquad V(\omega) \geqq 3^{\frac{1}{2}}/2.$$

Thus $\log|\eta(\Omega)|$ is approximated to within a bounded error by $\pi V(\omega)/12$, yielding, (after Dedekind [3]; p. 231),

$$(2.51)\qquad |\log \mathrm{H}(\omega) - \tfrac{1}{2}\log 2V(\omega) + \pi V(\omega)/6| \leqq \gamma_0,$$

and, easily,

$$(2.52)\qquad |\log \mathrm{H}(\omega)| \leqq \gamma_1 V(\omega),$$

where the γ_i will refer to positive absolute constants.

Now the estimation problem reduces easily from (2.11), and (2.52) to

$$(2.61)\qquad h\log\epsilon \leqq \gamma_2(\sum V(\omega_1) + \sum V(\omega_0)),$$

and finally to

$$(2.62)\qquad h\log\epsilon \leqq \gamma_2 \sum_{zy=k}\sum_{x=0} V([x+z\rho]/y),$$

dropping the restriction that y and $x+x\rho$ be relatively prime.

3. The $h\log\epsilon$ estimate.[1] We consider all the individual arguments of the "valence sum" (2.62), namely,

$$(3.11)\qquad \psi = (x+z\rho)/y = x/y + \alpha + \beta i.$$

with the abbreviations

$$(3.12)\qquad \alpha = z/(2y), \qquad \beta = 3^{\frac{1}{2}}z/(2y).$$

Consider a unimodular transformation in integral coefficients

$$(3.21)\qquad \Psi = (s'\psi - r')/(-s\psi + r), \qquad rs' - sr' = 1,$$

for which Ψ lies in the fundamental domain (2.42). Hence,

$$(3.31)\qquad 3^{\frac{1}{2}}/2 \leqq \operatorname{Im}\Psi = \beta\,|s(x/y+\alpha) - r + \beta i|^{-2} = V(\psi).$$

Therefore,

$$(3.32)\qquad s^2 \leqq 2\cdot 3^{-\frac{1}{2}}/\beta, \qquad\qquad (3.33)\qquad [r - s(x/y+\alpha)]^2 \leqq 2\beta/3^{\frac{1}{2}}.$$

when $s=0$ and (say) $r=1$, $V(\psi)=\beta$.

We finally rearrange the terms in valence sum (2.62) according to the

[1] The author is indebted to Professor A. Weil for suggestions leading to the simplification of the estimates.

fraction r/s (1/0 included). First of all $s=0$ only if $\beta \geqq 3^{\frac{1}{2}}/2$ or if $z \geqq y$. For such values of ψ, the valence sum is majorized by

$$(3.4) \qquad S_0 = \sum_{s=0} V([x+z\rho]/y) = (3^{\frac{1}{2}}/2) \sum_{zy=k} \sum_{x=0}^{y-1} (z/y) = 3^{\frac{1}{2}}\sigma(k)/2,$$

where $\sigma(k)$ represents the sum of the positive divisors of k.

Next we consider the portion of the valence sum restricted to fractions r/s with $s \neq 0$ and (say) positive. The contribution is written as

$$(3.51) \qquad S_1 = \sum_{zy=k}' \sum_{x=0}^{y-1}{}' \beta/\{[s(x/y+\alpha)-r]^2 + s^2\beta^2\},$$

where α, β always refer to abbreviations (3.12) and r/s is regarded as a function of x, y, and z (chosen earlier in equation (3.31)). The primes indicate that the sum is restricted so that $s > 0$.

The inside sum (formed by fixing y and z) can be estimated by further fixing s and regarding x as the variable with r dependent on x. The sum S_1 is amply majorized if we let x go from $-\infty$ to $+\infty$ for *each* r *that occurs.* We then can use the estimate

$$(3.52) \qquad \sum_{x=-\infty}^{\infty} \kappa/[\zeta^2 + (x-\xi)^2] \leqq \kappa\zeta^{-2} + \kappa \int_{-\infty}^{\infty} dx/(\zeta^2+x^2) = (\zeta^{-2} + \pi\zeta^{-1})\kappa$$

where $\zeta = \beta y$, $\xi = (ry/s) - y\alpha$, and $\kappa = \beta y^2/s^2$ in our context. In this manner the inside sum becomes transformed to a sum over *all* r/s *occurring* for a fixed y and z. The range of s is limited by inequality (3.32) to

$$(3.53) \qquad 1 \leqq s \leqq \mu, \qquad \mu^2 = 2 \cdot 3^{-\frac{1}{2}}/\beta = 4/3 \cdot y/z,$$

and for each s, the range of numerators r is limited by $4s$, since from inequality (3.33) combined with

$$(3.54) \qquad 2 \cdot 3^{-\frac{1}{2}} \geqq \beta = 3^{\frac{1}{2}}\alpha, \qquad 0 \leqq x < y,$$

we can conclude r lies in the interval $(-2/3^{\frac{1}{2}}, 5s/3 + 2/3^{\frac{1}{2}})$. We note finally from inequality (3.53), $2y \geqq 3z$, hence S_1 is majorized as follows:

$$(3.55) \qquad S_1 \leqq \sum_{\substack{zy=k\\ 2y \geqq 3z}} \sum_{s=1}^{\mu} 4s\beta y^2[\beta^{-2}y^{-2} + \pi/(\beta y)]/s^2,$$

$$(3.56) \qquad S_1 \leqq \sum_{\substack{zy=k\\ 2y \geqq 3z}} [2y/(3^{\frac{1}{2}}z) + \pi y] \sum_{1}^{\mu} 4/s, \qquad (3.57) \qquad S_1 \leqq \sum_{y|k} \gamma_3 y \log(4k/3),$$

since $\mu^2 = 4y/(3z) < 4k/3$. Hence the estimates (3.4) and (3.57) take us to

$$(3.6) \qquad h \log \epsilon \leqq \gamma_4 \sigma(k) \log k.$$

By the well known estimate [5], $\sigma(k) = O(k \log\log k)$, we obtain the result (1.1).

Perhaps a more concise way of presenting the basic result is to say that the average of the $9k''$ valences, $V((x+z\rho)/y)$, in (2.2), is $O(\log k)^{1+o(1)}$ as $k \to \infty$.

4. The class number estimate. To obtain an estimate for h from inequality (3.6) we consider the relation between ϵ and k using the following technique, (which must be presumed to be well-known although no specific reference comes to mind): Let $K'_{a,b}$ and $\bar{K}'_{a,b}$ be the conjugate fields to $K_{a,b}$ with ϵ' and $\bar{\epsilon}'$ corresponding to ϵ. Then since ϵ is a unit $|\epsilon'| = |\bar{\epsilon}'| = \epsilon^{-\frac{1}{2}}$, and Δ, the root-discriminant of ϵ, satisfies the inequality

$$(4.1) \qquad |\Delta| = |(\epsilon-\epsilon')^2(\epsilon-\bar{\epsilon}')^2(\epsilon'-\bar{\epsilon}')^2| \leqq (\epsilon+\epsilon^{-\frac{1}{2}})^4(2\epsilon^{-\frac{1}{2}})^2 \leqq \gamma_5\epsilon^3.$$

But $|\Delta| \geqq |d| = 3k^2$. Thus

$$(4.2) \qquad k \leqq \gamma_6\epsilon^{\frac{3}{2}},$$

which, together with estimate (3.6), yeilds

$$(4.3) \qquad h = O(\sigma(k))$$

and ultimately estimate (1.2). In both estimates (1.1) and (1.2) numerical constants could be easily supplied.

5. Norm estimate of modular invariants. A direct application of the sums (2.62) can be made to algebraic numbers of the type $j(\omega)$, where $j(z) = 1728J(z)$, for $J(z)$ the Klein modular function and ω a "positive-imaginary" quadratic number. For instance the quantities ω described in (2.2) are such that the $j(\omega)$ are a set of algebraic integers containing the conjugates of each element ([4]; p. 74). Thus taking one element, say $j(k\rho)$, we find

$$(5.1) \qquad \log|N(j(k\rho))| \leqq \gamma_7 \sum_{zy=k} \sum_{x=0}^{y-1} V([x+z\rho]/y)$$

and, as before,

$$(5.2) \qquad |N(j(k\rho))| \leqq \exp(\gamma_8\sigma(k)\log k)$$

where $N(\cdot\ \cdot\ \cdot)$ represents the (absolute) norm of an algebraic number.

In a more general situation the quadratic field will not have class number unity. Thus if λ is a positive-imaginary quadratic number, the

conjugates of $j(k\lambda)$ will be chosen from a finite number of sets of type (2.2), yielding ([4]; p. 76)

$$|N(j(k\lambda))| \leqq \exp(c(\lambda) k \log k \log\log k), \tag{5.3}$$

where $c(\lambda)$ is a constant depending on λ (and not on k).

NATIONAL BUREAU OF STANDARDS
AND
WASHINGTON UNIVERSITY.

REFERENCES.

[1] N. C. Ankeny, R. Brauer, and S. Chowla, "A note on the class numbers of algebraic number fields," *American Journal of Mathematics*, vol. 78 (1956), pp. 51-61.

[2] R. Brauer, "On the zeta-functions of algebraic number fields," *American Journal of Mathematics*, vol. 69 (1947), pp. 243-250.

[3] R. Dedekind, *Gesammelte mathematische Werke*, vol. II, Braunschweig, 1931.

[4] R. Fueter, *Vorlesungen uber die singulären Moduln und die komplexe Multiplikation der elliptischen Funktionen*, vol. I, Leipzig, 1924.

[5] G. H. Hardy, *An introduction to the theory of numbers*, Oxford, 1938, p. 264.

[6] H. Hasse, *Vorlesungen über Zahlentheorie, Berlin*, 1950, p. 396.

[7] E. Landau, "Abschätzung von Charaktersummen, Einheiten, und Klassenzahlen," *Nachrichten von der Königlichen Gesellschaft der Wissenschaften zu Göttingen*, 1918, p. 95.

[8] ———, "Verallgemeinerung eines Pólyaschen Satzes auf algebraische Zahlkörper," *ibid.*, 1918, p. 480.

FINITE FANO PLANES.*

By Andrew M. Gleason.

Introduction.[1] By a projective plane $\mathfrak{P}$ we will mean a set $\mathfrak{P}$ of points, certain subsets of which are called lines, such that the usual incidence axioms hold; viz.

I. If p and q are distinct points, there is a line L such that $p \in L$ and $q \in L$.

II. If L and M are distinct lines then $L \cap M$ contains exactly one point.

We shall also assume an axiom of non-triviality:

III. There exist four distinct points no three of which are members of the same straight line.

We definitely do not assume any further configuration axiom such as Desargues' without explicit mention. Although the notation is not symmetric, the axioms are well-known to be self-dual, so that any theorem about points and lines can be immediately reasserted after interchanging these concepts.

The cardinal of any two lines can easily be shown to be the same and this cardinal is the same as that of the set of lines containing a given point. If this cardinal is finite, say $n+1$, then n is called the order of the projective plane. In this case there will be n^2+n+1 points altogether and the same number of lines.

For any field F, or even a skew field, a model of this axiom system can be obtained by considering a three dimensional vector space over F, letting a "point" be a one-dimensional subspace and a "line" be a set of "points" which are subspaces of some two-dimensional subspace. In this model the classical configuration of Desargues will hold. Conversely, any projective plane in which Desargues' theorem is valid is isomorphic to one obtained by the preceding construction. One finds thus a one-to-one correspondence between isomorphism classes of Desarguesian projective planes and isomorphism classes of skew fields.

* Received February 15, 1956.

[1] The author is indebted to R. Baer, P. Dembowski, M. Hall, Jr., W. Pierce, and the referee for assistance in preparing the introduction and bibliography.

The analysis of projective planes with the aid of axioms I, II, and III only has not been very illuminating. A number of examples, some finite, have been constructed of projective planes in which Desargues' theorem fails. The finite ones have all been obtained by making systematic changes in the incidence structure of a Desarguesian plane; these changes do not affect the order of the plane which remains, in all known cases, a prime power. Whether this is a happenstance of our ignorance or a theorem has yet to be decided. Bruck and Ryser [2], by an ingenious number theoretic argument, have excluded an infinite list of integers as possible orders, and direct combinatorial study (finally using a giant computer [2]) has shown that only Desarguesian planes occur with orders less than 9.

More satisfactory results have been obtained by strengthening I, II, and III with other configuration theorems. Miss Moufang [8], for example, assumed that harmonic conjugates are uniquely determined by the usual construction; then, assuming also that no point is self-conjugate, she introduced affine coordinates from an alternative division ring. The small theorem of Desargues, taken as an axiom, also leads to coordinates from an alternative division ring (Hall [3]). Since a finite alternative division ring is necessarily a field (Zorn [13]), we can conclude that a finite projective plane satisfying the small theorem of Desargues is unrestrictedly Desarguesian.

The extra proviso of Miss Moufang that no point be self-conjugate leaves a gap in our knowledge. Assuming the unicity of the construction, if self-conjugacy occurs once it is universal [10]; this leads us to the study of planes in which this occurrence is postulated, which we call Fano planes. They are characterized by the fact that the diagonal points of any quadrangle are collinear.[3] Our main result is that a finite Fano plane is Desarguesian.

Assumptions on the nature of the collineation (automorphism) group are another way to supplement axioms I, II, and III. For a Desarguesian projective plane of order $n = p^m$ (where p is prime) the collineation group has order $mn^3(n-1)^2(n^2+n+1)(n+1)$. One expects, of course, that a non-Desarguesian plane will have far fewer collineations, and this is indeed the case so far as known. The collineation group of a Desarguesian plane always contains a cyclic transitive subgroup [12]; Hall [4] and Mann and Evans [7] have investigated the converse and have shown that, for $n \leq 1600$, the existence of a cyclic transitive group of collineations in a plane of order n

---

[2] Marshall Hall, personal communication.

[3] As remarked by the referee, there is a good deal to be said for calling such a plane an Anti-Fano plane, since the non-existence of a quadrilateral with collinear diagonal points was taken as an axiom by Fano.

implies that n is a prime power. In this paper we shall show that the existence in a finite plane of a relatively small number of collineations of a special type is sufficient to establish the small, and hence the full, theorem of Desargues; then we show that a finite Fano plane has sufficiently many of these special collineations.

Since our method makes particular use of the small Reidemeister configuration (called configuration G) it is closely related to the work of Klingenberg [5] who shows that the large Reidemeister configuration implies the theorem of Desargues.

For the sake of completeness we have included proofs of a number of already known results (see [9] for a comprehensive treatment of the subject), so that the paper is self-contained beyond the results of Hall just quoted on the small theorem of Desargues and the group theoretic interpretation of the small theorem given in Lemma 1.4.

1. Collineations. Consider a projective plane $\mathcal{P}$ satisfying axioms I, II, and III. A collineation of $\mathcal{P}$ is a permutation α of the set $\mathcal{P}$ such that $\alpha(x)$, $\alpha(y)$, and $\alpha(z)$ are collinear if and only if x, y, and z are collinear. Evidently the collineations form a group, Π. Any collineation of $\mathcal{P}$ induces a permutation of the set $\mathcal{L}$ of lines of $\mathcal{P}$, and this permutation preserves the phenomenon of concurrence. This new permutation representation of Π is evidently faithful. If every point of a certain line is left fixed by a collineation, we say the line is left fixed point-wise. Similarly, if every line through a point is fixed, we say that the point is left fixed line-wise.

For a given point x and line L consider the group G_{xL} of all collineations which leave x fixed line-wise and L fixed point-wise. In general we must expect that this group is trivial; that is, contains only the identity. We shall analyse the opposite possibility.

LEMMA 1.1. *If $\alpha \in G_{xL}$ and α leaves fixed a point y other than x and not on L (or, if α leaves fixed a line other than L and not passing through x), then α is the identity.*

Proof. Suppose that α leaves fixed y where $y \notin L$ and $y \neq x$. Then α leaves fixed every line M through y since y and $M \cap L$ are both fixed points. Therefore α leaves fixed every point which is the intersection of a line through x with a line through y. This includes all the points of $\mathcal{P}$ except those on the line joining x to y. Hence every line contains at least two fixed points and is itself fixed. Finally every point is fixed, being the intersection of two fixed lines. The other hypothesis is dual to the first.

LEMMA 1.2. *The set* $G_L = \bigcup_{x \in L} G_{xL}$ *is a group.* G_L *consists of the identity and those collineations of* $\mathfrak{P}$ *which leave* L *fixed pointwise, but have no other fixed points. If, for two distinct points* x *and* y *of* L, G_{xL} *and* G_{yL} *are non-trivial, then* G_L *is abelian; furthermore, every element of* G_L *(except the identity) has the same order, either infinity or a finite prime.*

Proof. By Lemma 1.1 any member of G_L fulfills the description given in the second statement. Conversely, consider a collineation α which leaves L pointwise fixed and has no other fixed points. For any $p \notin L$, let M be the line joining p and $\alpha(p)$; say M meets L at x. Now $\alpha(M)$ must contain both $\alpha(p)$ and $\alpha(x) = x$; so $\alpha(M) = M$ and M is fixed. Thus every point is on a fixed line. Any two of these fixed lines must meet at a fixed point which is therefore on L, so it must be x. Now we can see immediately that every line through x is fixed, so $\alpha \in G_{xL} \subset G_L$. This proves the second statement of the lemma.

Let α and β be any two elements of G_L, say $\alpha \in G_{xL}$ and $\beta \in G_{yL}$. To prove that G_L is a group, we must prove that $\alpha\beta^{-1}$ is in G_L. We may assume that neither α nor β is the identity and that $x \neq y$, since in these cases, $\alpha\beta^{-1} \in G_L$ trivially. Obviously, $\alpha\beta^{-1}$ is a collineation which leaves L pointwise fixed, so it will be sufficient to show that $\alpha\beta^{-1}$ has no fixed point not on L. Let p be any point not on L; set $q = \beta^{-1}(p)$ and $r = \alpha(q)$. Now $q, r = \alpha(q)$, and x are collinear and also $q, p = \beta(q)$, and y; since these two lines are distinct and meet at q we cannot have $p = \alpha\beta^{-1}(p) = r$ unless $p = q = r$; in this case, α and β would both have a fixed point not on L and would be the identity by 1.1. This case has been excluded.

We next prove that if x and y are distinct points of L, $\alpha \in G_{xL}$, and $\beta \in G_{yL}$, then α and β commute. To this end we must show that $\alpha\beta(p) = \beta\alpha(p)$ for any point p; this being trivial for $p \in L$, we assume the contrary. The following triples are collinear

$$\begin{array}{ll} x, p, \alpha(p) & y, p, \beta(p) \\ x, \beta(p), \alpha\beta(p) & y, \alpha(p), \beta\alpha(p) \\ x, \beta(p), \beta\alpha(p) & y, \alpha(p), \alpha\beta(p). \end{array}$$

(On the left, the first two derive from the fact that α leaves x fixed linewise; the third, is obtained from the first by applying β. Similarly on the right). These triples characterize both $\alpha\beta(p)$ and $\beta\alpha(p)$ as the point of intersection of the lines joining x to $\beta(p)$ and y to $\alpha(p)$. This proves that $\alpha\beta = \beta\alpha$.

Let us observe that if $x \neq y$, $G_{xL} \cap G_{yL} = \{\epsilon\}$, because any common

collineation satisfies the second hypothesis of Lemma 1.1. For the remainder of the proof we assume that at least two of the groups G_{xL} are non-trivial, which is equivalent to assuming that G_L contains each of the groups G_{xL} properly.

To prove that G_L is abelian, there is left only the case that α and β both come from the same subgroup G_{xL}. Choose a collineation $\gamma \in G_L - G_{xL}$; then certainly $\beta\gamma \notin G_{xL}$. By our previous result, α commutes with both γ and $\beta\gamma$; therefore, with β.

To demonstrate the last statement of the lemma, we suppose that G_L contains an element (other than ϵ) of finite order; this element can be chosen to have prime order p; let it be α and say $\alpha \in G_{xL}$. For any $\beta \notin G_{xL}$ say $\beta \in G_{yL}$ and $\alpha\beta \in G_{zL}$; here $y \neq z$. Now $\beta^p \in G_{yL}$ and also $\beta^p = \alpha^p\beta^p = (\alpha\beta)^p \in G_{zL}$; so $\beta^p = \epsilon$. Thus every element of $G_L - G_{xL}$ has order p; repeating the argument starting with β, we see that every element of G_L (except ϵ) has order p. This finishes the proof of Lemma 1.2.

LEMMA 1.3. *Suppose that* $\mathfrak{P}$ *is finite. If* L *is a line and* x *and* y *are two distinct points of* L *such that neither* G_{xL} *nor* G_{yL} *is trivial, then every element of either group (except* ϵ*) has the same prime order. If* x *is a point and* L *and* M *are two distinct lines through* x *such that neither* G_{xL} *nor* G_{xM} *is trivial, then every element of either group (except* ϵ*) has the same prime order.*

Proof. It is clear that, if $\mathfrak{P}$ is finite, G_L must be finite, so the first statement is part of the last lemma. The second statement is dual to the first. In applying duality, G_{xL} and G_{xM} must be interpreted as groups of permutations of the set $\mathfrak{L}$ of lines; but, since the statement concerns their abstract structure, it is immaterial.

LEMMA 1.4. *A necessary and sufficient condition that* G_L *be transitive on the points of* $\mathfrak{P} - L$ *is that the small theorem of Desargues hold on* L.

This has frequently been used in treatments of the coordinatization problem, so we omit the proof, which may be found, for example, in [1].

LEMMA 1.5. *Suppose that* $\mathfrak{P}$ *is finite of order* n. *The cardinal of* G_{xL} $(x \in L)$ *divides* n *and the cardinal of* G_L *divides* n^2. G_L *is transitive on* $\mathfrak{P} - L$ *if and only if the cardinal of* G_L *is exactly* n^2.

Proof. Let M be some line through x other than L. G_{xL} can be regarded as a group of permutations of the n points of $M - \{x\}$. Since no element of G_{xL} except the identity leaves fixed any point of $M - \{x\}$, each primitive

subset has the same cardinal as G_{xL}, and therefore this cardinal divides n. The other assertions follow in the same way when G_L is regarded as a group of permutations of the n^2 points of $\mathfrak{P} - L$.

Lemma 1.6. *Suppose that $\mathfrak{P}$ is finite of order n and that, for some line L, all of the groups G_{xL} $(x \in L)$ have the same cardinal $h > 1$. Then the small theorem of Desargues holds on L.*

Proof. Let g be the cardinal of G_L. We know that g divides n^2, say $gm = n^2$. In view of the last two lemmas we need only prove that $m = 1$. Now G_L is the union of $n + 1$ groups G_{xL}, each of cardinal h, with each pair having only the identity in common; hence, $(h-1)(n+1)+1 = g$, or $((h-1)(n+1)+1)m = gm = n^2$. This equation shows that $m < n$ (since $h > 1$) and $m \equiv 1 \pmod{n+1}$. Therefore, $m = 1$ and the lemma is proved.

Lemma 1.7. *Suppose that H is a group of permutations of a finite set L and suppose that, for some prime p and each $x \in L$, there exists an element of H of order p which leaves x fixed but has no other fixed points. Then H is transitive.*

Proof. Let L_1 be a primitive subset of L. Choose $x \in L_1$, and let α be an element of H of order p which leaves x fixed but has no other fixed points. L_1 consists of x and a number of disjoint α-primitive blocks each of cardinal p; hence, cardinal $L_1 \equiv 1 \pmod{p}$. Now suppose that H is not transitive, so that we can choose $y \in L - L_1$. Let β be an element of H which has order p and the unique fixed point y. Now L_1 is a union of disjoint β-primitive blocks each of cardinal p, hence cardinal $L_1 \equiv 0 \pmod{p}$. This contradiction establishes the lemma.

Theorem 1.8. *Let $\mathfrak{P}$ be a finite projective plane. Suppose, for every point x and line L with $x \in L$, that G_{xL} is non-trivial. Then $\mathfrak{P}$ is Desarguesian.*

Proof. Lemma 1.3 implies immediately that every non-trivial element of any of the groups G_{yM} $(y \in M)$ has the same prime order, say p.

Choose any line L and consider the group H_L of all collineations of $\mathfrak{P}$ which leave L fixed (not necessarily point-wise). H_L is represented (not faithfully) as a group of permutations of L. Now H_L contains as subgroups all of the groups G_{xM} where $x \in L$ and M is any line through x and a non-trivial element of G_{xM} $(M \neq L)$ has order p and acts on L with no fixed points except x itself; therefore, Lemma 1.7 applies and H_L is transitive on L. From the elementary formula $\alpha G_{xL} \alpha^{-1} = G_{\alpha(x) L}$ for $\alpha \in H_L$, it

now follows that all of the groups G_{xL} with $x \in L$ are conjugate in H_L. Now by Lemma 1.6, the small theorem of Desargues holds on the arbitrarily chosen line L. But a finite plane in which the small Desargues' theorem is valid is Desarguesian.

2. Projections. Consider a finite plane $\mathfrak{P}$ of order n, a definite line L of $\mathfrak{P}$ and a point x on L; we will obtain a sufficient conidtion that G_{xL} be non-trivial. Let $M_1, M_2, \cdots, M_n$ be an enumeration of the lines through x other than L. For any point a of $L - \{x\}$, let $\xi^a{}_{ji}$ denote the projection of M_i onto M_j from a; $\xi^a{}_{ji}$ is characterized by the fact that, for any $y \in M_i$, $\xi^a{}_{ji}(y) \in M_j$ and a, y, and $\xi^a{}_{ji}(y)$ are collinear. Obviously, $\xi^a{}_{kj}\xi^a{}_{ji} = \xi^a{}_{ki}$. If b is also a point of $L - \{x\}$, then $\xi^a{}_{ij}\xi^b{}_{ji}$ is a permutation of M_i. From the incidence axioms we deduce immediately that, for fixed a and b $(a \neq b)$, $J^{ab}{}_i = \{\xi^a{}_{ij}\xi^b{}_{ji} \mid j = 1, 2, \cdots, n\}$ is a simply transitive family of permutations of $M_i - \{x\}$.

LEMMA 2.1. *There is a one-to-one correspondence between elements of G_{xL} and permutations of M_i which commute with every permutation of the form $\xi^a{}_{ij}\xi^b{}_{ji}$. The correspondence is given by restricting elements of G_{xL} to M_i.*

Proof. Any element α of G_{xL} leaves fixed the line M_i; consequently, the restriction of α to M_i is a permutation of M_i. Obviously, the restriction map is a homomorphism of G_{xL}. Since no element of G_{xL} (except ϵ) can have any fixed points not on L, the restriction map is one-to-one.

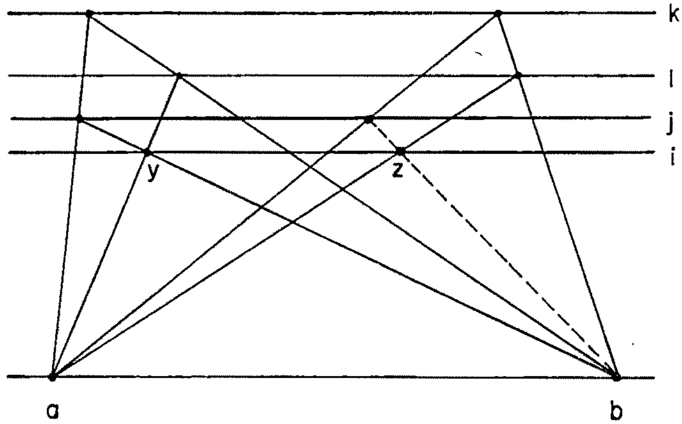
Suppose that $\alpha \in G_{xL}$, $y \in M_i$, $z \in M_j$, and $\xi^a{}_{ji}(y) = z$. We have $\alpha(a) = a$, $\alpha(y) \in M_i$, and $\alpha(z) \in M_j$ by the definition of G_{xL}; a, y, and z are collinear, so a, $\alpha(y)$, and $\alpha(z)$ are also; whence $\xi^a{}_{ji}\alpha(y) = \alpha(z) = \alpha\xi^a{}_{ij}(y)$. This being valid for any y on M_i, $\xi^a{}_{ji}\alpha = \alpha\xi^a{}_{ji}$. Then clearly, $\xi^a{}_{ij}\xi^b{}_{ji}\alpha = \alpha\xi^a{}_{ij}\xi^b{}_{ji}$; therefore α and its restriction to M_i commute with all permutations of the form $\xi^a{}_{ij}\xi^b{}_{ji}$.

Now suppose β is a permutation of M_i which commutes with all permutations of the form $\xi^a{}_{ij}\xi^b{}_{ji}$. If $a \neq b$ and $i \neq j$, the only fixed point of $\xi^a{}_{ij}\xi^b{}_{ji}$ is x, so β must leave x fixed; hence, we may define a permutation α of $\mathfrak{P}$ as follows: Pick a point a on $L - \{x\}$ and let $\alpha(y) = \xi^a{}_{ji}\beta\xi^a{}_{ij}(y)$ if $y \in M_j$ and $\alpha(y) = y$ if $y \in L$. We shall see that α is a collineation. To prove this it suffices to consider collinear points b, y, and z, where $b \in L$, $y \in M_j$, and $z \in M_k$, and show that b, $\alpha(y)$, and $\alpha(z)$ are collinear. We have $z = \xi^b{}_{kj}(y)$ and

$$\begin{aligned}\xi^b{}_{kj}\alpha(y) &= \xi^b{}_{kj}\xi^a{}_{ji}\beta\xi^a{}_{ij}(y) = \xi^a{}_{ki}(\xi^a{}_{ik}\xi^b{}_{ki})(\xi^b{}_{ij}\xi^a{}_{ji})\beta\xi^a{}_{ij}(y) \\ &= \xi^a{}_{ki}\beta(\xi^a{}_{ik}\xi^b{}_{ki})(\xi^b{}_{ij}\xi^a{}_{ji})\xi^a{}_{ij}(y) = (\xi^a{}_{ki}\beta\xi^a{}_{ik})\xi^b{}_{kj}(y) = \alpha(z);\end{aligned}$$

that is, b, $\alpha(y)$, and $\alpha(z)$ are collinear. It is clear that α leaves L fixed point-wise and x fixed line-wise; i.e., $\alpha \in G_{x\,L}$. This completes the proof of the lemma.

We will now assume that, for any a and b of $L - \{x\}$ and any indices i, j, k, l, if $\xi^b{}_{ij}\xi^a{}_{jk}\xi^b{}_{kl}\xi^a{}_{li}$ (which is a permutation of M_i) has a fixed point, then it is the identity. This corresponds to the closure of the following figure. Here we have taken x at infinity. The diagram shows y as a fixed point of the fourfold projection and the closure of the figure shows that another point z is also left fixed. The figure consists of 13 lines and 11 points



(counting x); five lines go through each of the three points a, b, and x, and three through every other point. (There is a degenerate case if M_j coincides with M_l or M_i with M_k. The closure of the degenerate case is implied by the closure of all non-degenerate cases.) This configuration will be referred to as configuration G because it implies that the set $J^{ab}{}_i$ is a group. We shall say that the configuration G holds in $\mathfrak{P}$ if it holds regardless of the choice of x and L. It is easy to show that configuration G holds in any Desarguesian plane. This configuration has been studied by Reidemeister [11] who obtains the following lemma.

LEMMA 2.2. *If configuration G holds in $\mathfrak{P}$ then the set $J^{ab}{}_i$ is a group. Conversely, if $J^{ab}{}_i$ is a group for all choices of L, x, a, b, and i, then configuration G holds in $\mathfrak{P}$.*

Proof. Suppose that $\xi^a{}_{ik}\xi^b{}_{ki}$ and $\xi^a{}_{il}\xi^b{}_{li}$ are two given members of $J^{ab}{}_i$. Choose any point y of M_i and an index j such that $(\xi^a{}_{ik}\xi^b{}_{ki})(\xi^a{}_{il}\xi^b{}_{li})^{-1}(y) = \xi^a{}_{ij}\xi^b{}_{ji}(y)$. This is possible since $J^{ab}{}_i$ is a transitive family. Then y is a fixed point of $\xi^b{}_{ij}\xi^a{}_{ji}\xi^a{}_{ik}\xi^b{}_{ki}\xi^b{}_{il}\xi^a{}_{li} = \xi^b{}_{ij}\xi^a{}_{jk}\xi^b{}_{kl}\xi^a{}_{li}$; hence this permutation

has only fixed points. It follows immediately that $(\xi^a{}_{ik}\xi^b{}_{ki})(\xi^a{}_{il}\xi^b{}_{li})^{-1} = \xi^a{}_{ij}\xi^b{}_{ji}$; thus $J^{ab}{}_i$ is a group.

Conversely, assume that $J^{ab}{}_i$ is always a group. If $\xi^b{}_{ij}\xi^a{}_{jk}\xi^b{}_{kl}\xi^a{}_{li} = (\xi^a{}_{ij}\xi^b{}_{ji})^{-1}(\xi^a{}_{ik}\xi^b{}_{ki})(\xi^a{}_{il}\xi^b{}_{li})^{-1}$ has a fixed point, it is the identity, being an element of $J^{ab}{}_i$ which is always a simply transitive family of permutations.

LEMMA 2.3. *Suppose that A, B, and C are subgroups of a group and suppose that the elements of A and B can be so indexed that, for each i, $\alpha_i\beta_i \in C$. Assume, moreover, that every element of C can be written in the form $\alpha_i\beta_i$. Then A and B commute as complexes; i. e., $AB = BA$.*

Proof. Given α_i and β_j, choose k so that $(\alpha_j\beta_j)(\alpha_i\beta_i) = \alpha_k\beta_k$; then $\beta_j\alpha_i = (\alpha_j^{-1}\alpha_k)(\beta_k\beta_i^{-1})$; proving $BA \subset AB$. Taking inverses, the latter relation becomes $AB \subset BA$, whence $AB = BA$.

LEMMA 2.4. *If $A_1, A_2, \cdots, A_n$ is a sequence of subgroups of a group which mutually commute as complexes, then $B = A_1A_2\cdots A_n$ is a group. Moreover, if the groups A_i are finite, then B is finite and every prime factor of the cardinal of B is a factor of the cardinal of A_i, for some i.*

Proof. This is well-known for $n = 2$. The general case follows easily by induction.

THEOREM 2.5. *If configuration G holds in $\mathfrak{P}$ and the order of $\mathfrak{P}$ is a prime power, then $\mathfrak{P}$ is Desarguesian.*

Proof. Let L be a fixed line of $\mathfrak{P}$ and x, a, b, and c points of L. Observe that $J^{ab}{}_i = J^{ba}{}_i$ because the latter set is made up of the inverses of the members of the former, which is a group. Enumerating $J^{ab}{}_i$ in the order $\xi^b{}_{ij}\xi^a{}_{ji}$, $j = 1, 2, \cdots, n$ and $J^{ac}{}_i$, $\xi^a{}_{ij}\xi^c{}_{ji}$, the corresponding products $\xi^b{}_{ij}\xi^a{}_{ji}\xi^a{}_{ij}\xi^c{}_{ji} = \xi^b{}_{ij}\xi^c{}_{ji}$ run through the group $J^{bc}{}_i$; therefore, the groups $J^{ab}{}_i$ and $J^{ac}{}_i$ commute as complexes. Since the groups $J^{ab}{}_i, J^{ac}{}_i, \cdots, J^{ae}{}_i$, where $b, c, \cdots, e$ runs through all the points of $L - \{x\}$, mutually commute, the product $K = J^{ab}{}_iJ^{ac}{}_i\cdots J^{ae}{}_i$ is a group. Each of the groups $J^{ad}{}_i$ has cardinal $n = p^m$, where p is a prime, so the only prime which divides the cardinal of K is p. Thus K is a p-group, and therefore K contains a non-trivial central element β. Since $J^{ab}{}_iJ^{ac}{}_i \supset J^{bc}{}_i$ (as we showed above) we know that K contains all groups of the form $J^{bc}{}_i$, a fortiori, all permutations of the form $\xi^b{}_{ij}\xi^c{}_{ji}$; so β commutes with all these permutations. By Lemma 2.1, G_{xL} is non-trivial. Here x and L were chosen arbitrarily except for the condition $x \in L$, so by Theorem 1.8, $\mathfrak{P}$ is Desarguesian.

3. Fano Planes. We shall call a projective plane a Fano plane if the diagonal points of every quadrangle are collinear.

LEMMA 3.1. *In a Fano plane every permutation of the form* $\xi^a{}_{ij}\xi^b{}_{ji}$ $(a \neq b, i \neq j)$ *has order two. Hence it satisfies* $\xi^a{}_{ij}\xi^b{}_{ji} = \xi^b{}_{ij}\xi^a{}_{ji}$.

Proof. Take any point y on M_i and consider the quadrangle with vertices x, b, $\xi^b{}_{ji}(y)$, and $\xi^a{}_{ij}\xi^b{}_{ji}(y)$. Its diagonal points are a, y, and $\xi^b{}_{ji}\xi^a{}_{ij}\xi^b{}_{ji}(y)$, whence $y = \xi^a{}_{ij}\xi^b{}_{ji}\xi^a{}_{ij}\xi^b{}_{ji}(y)$. Since y is arbitrary $(\xi^a{}_{ij}\xi^b{}_{ji})^2 = \epsilon$.

LEMMA 3.2. *In a Fano plane, the permutations* $\xi^a{}_{ij}\xi^b{}_{ji}$ *and* $\xi^a{}_{ik}\xi^b{}_{ki}$ *commute.*

Proof. In any group, if α, β, and $\alpha\beta$ are of order two, then α and β commute; hence we will prove that $((\xi^a{}_{ij}\xi^b{}_{ji})(\xi^a{}_{ik}\xi^b{}_{ki}))^2 = \epsilon$. Using Lemma 3.1,

$$((\xi^a{}_{ij}\xi^b{}_{ji})(\xi^a{}_{ik}\xi^b{}_{ki}))^2 = (\xi^a{}_{ij}\xi^b{}_{ji})(\xi^b{}_{ik}\xi^a{}_{ki})(\xi^a{}_{ij}\xi^b{}_{ji})(\xi^b{}_{ik}\xi^a{}_{ki})$$
$$= \xi^a{}_{ij}(\xi^b{}_{jk}\xi^a{}_{kj})^2\xi^a{}_{ji} = \xi^a{}_{ij}\xi^a{}_{ji} = \epsilon.$$

LEMMA 3.3. *If a permutation group is abelian and transitive, then it is simply transitive.*

Proof. To prove that there is exactly one element which effects any given partial map $x \to y$, it is sufficient to consider only the case $y = x$ and prove that the only element leaving x fixed is the identity. Suppose α leaves x fixed. For any z we can choose β so that $\beta(z) = x$. Then $\alpha(z) = \beta^{-1}\alpha\beta(z) = z$; so α leaves every point fixed; i. e., α is the identity.

LEMMA 3.4. *In a Fano plane, the sets* $J^{ab}{}_i$ *are elementary abelian 2-groups.*

Proof. We have already seen that $J^{ab}{}_i$ is a set of mutually commuting permutations of $M_i - \{x\}$ which is transitive on $M_i - \{x\}$. Therefore the group generated by $J^{ab}{}_i$ is an abelian transitive permutation group. But such a group is simply transitive, therefore $J^{ab}{}_i$ must be the entire group. Since every element of $J^{ab}{}_i$ (except ϵ) has order two, it is an elementary abelian 2-group.

THEOREM 3.5. *A finite Fano plane is Desarguesian.*

Proof. We know that the sets $J^{ab}{}_i$ are always 2-groups, therefore configuration G holds. If $\mathfrak{P}$ is finite, then the order of $\mathfrak{P}$ is the cardinal of $J^{ab}{}_i$ which is a power of two, so the theorem follows directly from Theorem 2.5.

HARVARD UNIVERSITY.

REFERENCES.

[1] E. Artin, *Selected topics in geometry* (mimeographed), New York University, 1954.
[2] R. Bruck and H. Ryser, "The nonexistence of certain finite projective planes," *Canadian Journal of Mathematics*, vol. 1 (1949), pp. 88-93.
[3] Marshall Hall, Jr., "Projective planes," *Transactions of the American Mathematical Society*, vol. 54 (1943), pp. 229-277.
[4] ———, "Cyclic projective planes," *Duke Mathematical Journal*, vol. 14 (1947), pp. 1079-1090.
[5] W. Klingenberg, "Beweis des Desarguesschen Satzes aus der Reidemeisterfigur und verwandte Sätze, *Abhandlungen aus dem Mathematischen Seminar der Hamburgischen Universität*, vol. 19 (1955), pp. 158-175.
[6] L. Lombardo-Radici, "Sui piani microdesarguesiani affini," *Atti della Accademia Nazionale Lincei Rendiconti Classe Scienze fisiche Matematiche Naturali*, (8), vol. 15 (1953), pp. 264-271.
[7] H. B. Mann and T. A. Evans, "On simple difference sets," *Sankhya*, vol. 11 (1951), pp. 357-364.
[8] R. Moufang, "Alternativkorper und der Satz von vollständigen Viereck," *Abhandlungen aus dem Mathematischen Seminar der Hamburgischen Universität*, vol. 9 (1933), pp. 207-222.
[9] G. Pickert, *Projektive Ebene*, Springer, Berlin, 1955.
[10] ———, "Der Satz vom vollständigen Viereck bei kollinearen Diagonalpunkten," *Mathematische Zeitschrift*, vol. 56 (1952), pp. 131-133.
[11] K. Reidemeister, "Topologische Fragen der Differentialgeometrie V," *Mathematische Zeitschrift*, vol. 29 (1929), pp. 427-435.
[12] J. Singer, "A theorem in finite projective geometry and some applications to number theory," *Transactions of the American Mathematical Society*, vol. 43 (1938), pp. 377-385.
[13] M. Zorn, "Theorie der alternativen Ringe," *Abhandlungen aus dem Mathematischen Seminar der Hamburgischen Universität*, vol. 8 (1931), pp. 123-147.

CONTRIBUTION TO THE PICARD-VESSIOT THEORY OF HOMOGENEOUS LINEAR DIFFERENTIAL EQUATIONS.*

By A. SEIDENBERG.

1. Introduction. Let F be an ordinary differential field of characteristic 0, C its field of constants, and $L(y) \equiv y^{(n)} + p_1 y^{(n-1)} + \cdots + p_n y = 0$, $p_i \in F$, a linear homogeneous differential equation of order n. Let η be a non-trivial solution of this equation and let D be the field of constants of $F\langle\eta\rangle$. One question in the Picard-Vessiot theory is whether η can be so selected that $D = C$. A general result of E. R. Kolchin [3], which is not restricted to linear equations, shows that a solution exists for which D is algebraic over C; and hence disposes of the question if C is algebraically closed. If C is not necessarily algebraically closed, then according to a result of M. P. Epstein [2], one can at any rate choose a fundamental system of solutions $\eta_1, \cdots, \eta_n$ of $L(y)$ so that the field of constants E of $F\langle\eta_1, \cdots, \eta_n\rangle$ is normal over C. The conjecture, however, that one can always arrange to have $D = C$ (hence also $E = C$) turns out to be false, even for $n = 2$, as we show by a counter-example below. Although placed last, this counter-example can be read separately and in itself answers the conjecture. It seems preferable, however, to have a necessary and sufficient condition that $D = C$ be obtainable. This is done below for $n = 2$. This throws light on the counter-example and also prepares the way, possibly, for consideration of the case $n > 2$.

2. Reduction to a first order equation. For a first order equation $y' = py$ one can obtain $D = C$ without difficulty and a proof can be found in [2]. One may show, in fact, (by an argument occurring in Section 3, below) that if the general solution[1] y of $y' = py$ introduces a constant, then

* Received July 5, 1956.

[1] By a *solution* to a differential equation $G(Y, Y_1, \cdots, Y_r) = 0$, where $G \in F\{Y\} = F[Y, Y_1, \cdots]$ we mean a quantity y in a differential extension field of the base field F such that $G(y, y', \cdots, y^{(r)}) = 0$. If $G(Y)$ is an irreducible polynomial involving Y_r, then there is a solution y such that degree of transcendence of $F\langle y\rangle/F = r$; and $F\langle y\rangle/F$ is uniquely determined up to an isomorphism by this condition: we refer to this solution as the *general solution* of $G(Y) = 0$. Caution: it is not in general true that every solution of $G(Y) = 0$ is a specialization of the general solution: for illustrative examples, see J. F. Ritt's Colloquium Series book, *Differential Algebra*,

CAUCHY'S STABLE DISTRIBUTIONS AND AN "EXPLICIT FORMULA" OF MELLIN.*

By Aurel Wintner.

1. If a (or its real part) is positive and less than 1, and if t (or its real part) is positive, then, according to Mellin, the case $\phi(x) = \exp(-tx^a)$ of the Euler-MacLaurin difference $\Sigma\phi(n) - \int\phi(x)\,dx$, where $n = 1, 2, \cdots$ and $0 \leqq x < \infty$, can be evaluated explicitly; cf. (9) and (9 bis) below. The resulting "explicit formula" turns out to be closely connected with results which in earlier papers I obtained on Cauchy's symmetric stable distribution functions of index a (where either $0 < a < 1$, as in Mellin's formula, or, to certain ends, $0 < a < 2$, as in Paul Lévy's justification of Cauchy's formal result). This family of distribution functions is defined by the property that the Fourier transform of the distribution is even and is identical with e^{-t^a} if $0 \leqq t < \infty$.

In what follows, the analytical questions concerning Cauchy's distributions will be dealt with in a systematic way. Roughly speaking, there will be three issues involved: (i) the explicit form of the Laplace transform of the distribution, (ii) the connection of the Laplace transform with the Stieltjes transform, finally (iii) the application of the resulting explicit formulae to a determination of the weight function which belongs to the "stratification," in terms of the symmetric Gaussian distribution ($\beta = 2$), of the symmetric stable distribution of any index $\beta < 2$.

At the end (Section 13), there will be considered the asymptotic expansion (and its implications for Cauchy's transcendents) which takes the place of Mellin's convergent expansion if $0 < a < 1$ is replaced by $1 < a < \infty$ (the limiting case $a = 1$ is trivial). The values $a = 2n$, where $n = 2, 3, 4, \cdots$, are exceptional (but $n = 1$, i.e., $a = 2$, is trivial). For this exceptional case, there will be reproduced with Professor Pólya's permission an asymptotic formula which he communicated to me several years ago.

For $0 < a < 2$, a side issue will be the "shape" of the frequency curves. This issue has, however, little to do with Cauchy's particular choice of symmetric distributions, since what results holds for all symmetric distributions

* Received June 11, 1956.

of the so-called L-class (Khintchine-Lévy). This and related results are dealt with in Appendix II and Appendix III.

Appendix I deals again with the entire range $0 < a < \infty$ but in the angular case. What is then involved is the extension from the case $a = 2$ of an elliptic theta to the case of an arbitrary $a > 0$ (including the exceptional a-values $4, 6, \cdots$) of the function

$$1 + 2 \sum_{n=1}^{\infty} q^{n^a} \cos nx.$$

This function has of course often been considered. The results of Appendix I are however substantially sharper than those obtained in the literature consulted.

2. Replace the function e^{-t^a}, where $0 \leqq t < \infty$, by its Fourier cosine transform

$$F(t) = \int_0^{\infty} e^{-s^a} \cos ts \, ds. \tag{1}$$

Then a partial integration shows that

$$F(t) = G(1/t^a)/t^{1+a}, \tag{2}$$

where

$$G(t) = \int_0^{\infty} e^{-ts} \sin(s^{1/a}) \, ds \tag{3}$$

(even if only just $0 < a < \infty$, rather than $0 < a < 1$, is assumed).

The representation (2) of (1) in terms of the Laplace transform (3) was pointed out in [13], p. 678. The correspondence (2) between F and G is invariant under the substitution

$$(F, t, a) \rightarrow (G, 1/t, 1/a) \tag{2*}$$

(if $0 < t < \infty$ and $0 < a < \infty$), since (2) is equivalent to

$$G(t) = F(1/t^{1/a})/t^{1+1/a}. \tag{2 bis}$$

The alternative devices, $F \rightarrow G$ and $G \rightarrow F$, can be applied to each of the Fourier analyses to be considered.

An immediate consequence of (2) and (3) is the following asymptotic relation of Pólya (cf. Pólya-Szegö, chap. III, no. 154):

$$F(t) \sim A t^{-1-a} \text{ as } t \rightarrow \infty, \text{ where } A = \Gamma(1 + a) \sin(\tfrac{1}{2}\pi a) \tag{4}$$

(and where the $\sim$ is meant in the sense that

(4 bis) $$F(t) = o(t^{-1-\alpha}) \text{ as } t \to \infty \text{ if } \alpha = 2, 4, \cdots,$$

the constant A being 0 when α is an even integer). In fact, it is well-known that

$$\int_0^\infty \sin(s^{1/\alpha})\,ds = \Gamma(1+\alpha)\sin(\tfrac{1}{2}\pi\alpha),$$

provided that integral on the right, which is convergent only if $0 < \alpha < 1$, is meant as a Cesàro limit (of an appropriate positive index) when $1 \leqq \alpha < \infty$. Since summability in Abel's sense is implied by (and is a process consistent with) summability in Cesàro's sense, (4) follows by using (2) for $t \to \infty$ and letting $t \to +0$ (Abel) in (3).

It is clear from (1) that $F(t) > 0$ holds for $t = 0$, hence for all sufficiently small $t > 0$, while (4) shows that, if $2n < \alpha < 2(n+1)$, where $n = 0, 1, \cdots$, then $F(t) > 0$ or $F(t) < 0$ holds for all sufficiently large t according as n is even or odd. (In view of (4 bis), nothing follows for large t if $\alpha = 2, 4, \cdots$, but Pólya has shown that (1) is then an entire function having no zero at all or only real zeros according as $\alpha = 2$ or $\alpha = 4, 6, \cdots$; cf. Pólya-Szegö, chap. V, no. 170).

3. Let $\psi(s)$, where $0 \leqq s < \infty$, be a real-valued (not necessarily continuous) function whose Laplace transform

(5) $$\int_0^\infty e^{-ts}\psi(s)\,ds$$

is convergent for $t > 0$. Let N, where $0 \leqq N \leqq \infty$, denote the number of the positive zeros t of (5), and let M, where $0 \leqq M \leqq \infty$, denote the number of those s-values at which the function

(5 bis) $$\Psi(s) = \int_0^s \psi(u)\,du$$

changes sign when s increases from $s = +0$ to ∞. Then $N \leqq M$. This inequality expresses a theorem of Laguerre as further developed by Pólya (cf. Pólya-Szegö, chap. V, no. 82).

Due to (2) and (3), the estimate $N \leqq M$ can be applied to (1), with

(6) $$\psi(s) = \int_0^s \sin(u^{1/\alpha})\,du$$

(so that

$$\Psi(s)=\int_0^{\varepsilon}(s-u)\sin(u^{1/a})\,du, \tag{6 bis}$$

by (5 bis)). In fact, if $N=N_a$ and $M=M_a$ denote the respective number of times the functions (1) and (6 bis) change sign when t and s vary from $+0$ to ∞, then it follows that

$$0\leqq N_a\leqq M_a\leqq\infty,\text{ where } 0<a<\infty.$$

In fact, it is clear from (2) that it is immaterial whether N_a is referred to $F(t)=F_a(t)$ or to $G(t)=G_a(t)$. On the other hand, a partial integration of (3) shows that $G(t)/t$ is identical with (5) if $\psi(s)=\psi_a(s)$ is defined by (6). This proves the assertion of the last formula line.

Another straightforward consequence (which is related to, but does not involve, the inequality $N\leqq M$) is the complete monotony of $G(t)t=G_a(t)/t$ on $(0,\infty)$ if $0<a\leqq 1$ (the limiting case $a=1$ is trivial, since (2) can be evaluated explicitly if $a=1$). In fact, the substitution $u^{1/a}=v$ shows that the function (6) is non-negative for $0\leqq s<\infty$ if $0<a<1$ (simply because v^{a-1} is positive and non-increasing for $0<v<\infty$ if $0<a<1$). But if $\psi\geqq 0$ throughout, then the function (5) is completely monotone on the half-line $0<t<\infty$. Since $G(t)/t$ is identical with the case (6) of (5), the assertion follows.

A corollary of this fact is that the function

$$F(1/t)/t^{1+2a},\text{ where } 0<t<\infty, \qquad (F=F_a),$$

possesses a Hausdorff-Bernstein representation if $0<a<1$ (and, by continuity, also in the limiting case $a=1$ which, however, is a trivial case). In fact, if $0<t<\infty$, $0<a<1$ and $m=1,2,\cdots$, then the m-th derivative of t^a is positive or negative according as m is odd or even. Hence, successive differentiations show that if a function $g(t)$ is completely monotone on $(0,\infty)$, then the same is true of the function $g(t^a)$, where $0<a<1$. Hence, the assertion follows by choosing $g(t)=G(t)/t$, since (2) shows that $G(t^a)/t^a$ is identical with the function occurring in the last formula line. (For an application of the substitution $t\to t^a$ to $g(t)=e^{-t}$, instead of the more elaborate function $g(t)=G(t)/t$, cf. (19)-(21) below.)

If $a=1$, then (1) and (3) show that both $F(t)$ and $G(t)$ reduce to $(1+t^2)^{-1}$ (which argrees with, though it is not implied by, (2) if $a=1$). Hence the limiting case, $a=1$, of the preceding result on $G(t)/t$ means that $(t+t^3)^{-1}$ is completely monotone on $(0,\infty)$. The same cannot be true of

$G(t)$ itself, since already the first derivative of $(1+t^2)^{-1}$ fails to be of constant sign on $(0,\infty)$.

5. If t is complex, then the integral (3) is absolutely convergent to the right, and divergent to the left, of the imaginary axis of the t-plane, whenever $0<\alpha<\infty$. But if the point $t=0$ is excluded when $\alpha\geqq 1$, then analytic continuation of $G(t)$ is possible and does not lead to any singularity, and the exclusion of point $t=0$ is unnecessary when $0<\alpha<1$. In fact, it was shown in [13] that if $0<\alpha<1$, then $G(t)$ is a transcendental entire function, having MacLaurin expansion

$$(7)\qquad G(t)=\sum_{n=1}^{\infty} c_{n-1}t^{n-1}/\Gamma(n),\qquad c_{n-1}=(-1)^{n-1}\alpha\Gamma(\alpha n)\sin(\tfrac{1}{2}\pi\alpha n).$$

The radius of convergence of the power series (7) is ∞, 1 or 0 according as $0<\alpha<1$, $\alpha=1$ or $1<\alpha<\infty$. In the third case, the series (7) is an asymptotic expansion of $G(t)$ for $t\to+0$ which, in view of (2), supplies an asymptotic expansion of $F(t)$ for $t\to+\infty$. In the second case, (7) reduces to $G(t)=(1+t^2)^{-1}$, since c_{n-1} becomes $(-1)^{n-1}\Gamma(n)(-1)^n=-\Gamma(n)$ or 0 according as n is odd or even. In all three cases, the functional equation

$$(8)\qquad \Gamma(1+z)\Gamma(1-z)=\pi z/\sin\pi z$$

leads to some reduction of (7).

As pointed out at the end of [13], the function (3) and its expansion (7) are "formally related to the standard entire functions occurring in the theories of explicit analytic continuations beyond the circle of convergence of a power series," namely, to Mittag-Leffler's transcendents E. But neither (1)-(3) nor (7) is mentioned in the subsequent literature, listed in [1], where [13] is overlooked, as is [14]. Cf. also the footnote at the end of Section 7 below.

6. An approach to (1)-(3) and (7) which is different from that followed in [13] was found in [14], pp. 705-707; it was shown there that the true source of the expansion (7) of (3) is a formula of Mellin [8], p. 12 (which I did not find quoted in the textbooks consulted or in the relevant papers, such as the fundamental paper of Hardy and Littlewood [3], p. 136; cf. p. 120, footnote, which refers to Mellin's work as a whole, but not to his paper [8]; cf. also Lindelöf's general theory in [7]). The formula in

question is an "explicit formula" (in the sense of the term customary in the analytic theory of numbers); it states that

$$(9)\qquad t^{1/a}\sum_{k=1}^{\infty} e^{-tk^a} = \Gamma(1+1/a) + \sum_{n=1}^{\infty}(-1)^n\zeta(-an)t^n/\Gamma(n+1),$$

where the numerical value of the leading term on the right is produced, of course, by the trivial identity

$$(9\text{ bis})\qquad t^{1/a}\int_0^{\infty} e^{-tx^a}\,dx = \Gamma(1+1/a),$$

and where (in order to assure convergence in the entire t-plane, rather than just the validity of (9) as an asymptotic expansion as $t\to+0$) it is assumed that $0<a<1$, rather than only that $0<a\neq 2,4,\cdots$. As shown in [14], pp. 705-707, the expansions (7) and (9) are equivalent by virtue of the functional equations (8) and

$$(10)\qquad \eta(z)=\eta(1-z),\ \text{where}\ \eta(z)=\pi^{-\frac{1}{2}z}\Gamma(\tfrac{1}{2}z)\zeta(z),$$

and of a Möbius inversion.

Actually, (9) leads not only to the expansion (7) of $G(t)$ but also to the Laplace analysis of $F(t)$, exhibited by (12) (and (13), (11)) below. In fact, (12) below and (7) are equivalent by virtue of (1)-(3), as seen by a term-by-term integration (the legitimacy of which is trivial from $0<a<1$). This is precisely the route along which the expansion (7) was obtained in [13].

7. Let $0\leqq t<\infty$. Then the Laplace analysis (that is, the determination of the *Unterfunktion*) of $G(t)$ is given by (3), where only $a>0$ is asumed, and, as mentioned above, (3) follows from the two definitions (1), (2) by nothing deeper than a partial integration. In contrast, the Laplace analysis of $F(t)$ results only after a contour integration, which succeedss directly only if

$$(11)\qquad 0<a<1.$$

The result is

$$(12)\qquad F(t)=\int_0^{\infty} e^{-ts}\exp(-\lambda s^a)\sin(\mu s^a)\,ds,\qquad (0\leqq t<\infty)$$

where $\lambda=\lambda(a)$, $\mu=\mu(a)$ are abbreviations for the constants

$$(13)\qquad \lambda=\cos\tfrac{1}{2}\pi a,\qquad \mu=\sin\tfrac{1}{2}\pi a,\qquad (\lambda^2+\mu^2=1).$$

It is clear from (11) that $t = 0$ is the abscissa of convergence, as well as the abscissa of absolute convergence, of the Laplace transform (12). Incidentally, (11) also implies that both constants (13) are positive.

The Laplace analysis (12) of $F(t)$ is known; in fact, it is contained in what was obtained in [12], pp. 86-88, for the m-dimensional generalization (from $m = 1$) of the "symmetric stable frequency function" of Cauchy which, if $m = 1$, becomes $1/\pi$ times the function (1) itself (it may be mentioned that subsequently Lévy's monograph [6], pp. 221-224, and a dissertation written under his direction (cf. p. 221, footnote) have transferred his fundamental enumeration (1923) of all, not necessarily symmetric, stable frequency functions from $m = 1$ to any m; but the Laplace analysis of these functions is not taken up *loc. cit.*, not even in the symmetric case).

The proof of (12) can be sketched as follows (for details, cf. [12], pp. 86-88): In view of (1), where t can be thought of as a fixed positive number, it is sufficient to deal with the real part of the integral

$$(14) \qquad \int_0^\infty u^{-1+1/\alpha} \exp(-u + itu^{1/\alpha})\, du,$$

where $u = s^{1/\alpha} \geqq 0$. If $L(\phi)$ denotes, for a fixed ϕ, the half-line $\arg u = \phi$ in the complex u-plane, then $L(0)$ is the path of integration in (14). But no singularity of the integrand is met when $L(\phi)$ sweeps through the wedge $0 \leqq \phi \leqq \frac{1}{2}\pi\alpha$ $(< \frac{1}{2}\pi)$, except for the common end-point, $u = 0$, which is a harmless singularity (in fact, $-1 + 1/\alpha > 0$). Hence, if $L(0)$ is deformed into $L(\frac{1}{2}\pi\alpha)$, then (12) follows from (11) and (13), simply by taking the real part of the $L(\frac{1}{2}\pi\alpha)$-representation of (14).*

8. From the Laplace analysis, (12), of the Fourier cosine transform, (1), of e^{-t^α}, where $0 \leqq t < \infty$, the Stieltjes analysis of e^{-t^α} itself can be

* According to [1], it was conjectured by W. Feller and proved by H. Pollard that $E(-t)$, where $E(z) = E_\alpha(z)$ denotes Mittag-Leffler's entire function $\sum z^n/\Gamma(\alpha n + 1)$, is completely monotone on the half-line $0 \leqq t < \infty$ if the index α is on the range (11) (this is true, but trivial, in the limiting cases, $\alpha = 0$ and $\alpha = 1$, of (11), since $E_0(-t) = (1 + t^2)^{-1}$ and $E_1(-t) = e^{-t}$). What is involved here are two papers of Pollard [10] which appeared as late as 1946 and 1948, and which will be referred to as (i) and (ii) respectively. Both (i) and (ii) overlook [12], pp. 86-88, and [13] (cf. also [14], pp. 705-707). But the content of (i) is just an artificial verification of a wrong form of (6) and/or (7) above (in order to see that there is a mistake somewhere, it is sufficient to choose Pollard's λ, which is the α in (11) above, to be $\frac{1}{2}$ in formulae (5) and (4) of (i), which correspond to (7) and (12) above). Correspondingly, the complete monotony of Mittag-Leffler's $E(-t)$ is concluded in (ii) from erroneous formulae. The necessary corrections can be read off from [13] and/or [14].

obtained. In other words, it is possible to obtain the explicit form of the function $h(s)$ for which

$$(15) \qquad e^{-t^a} = \int_0^\infty (t^2+s^2)^{-1} h(s)\, ds$$

holds as an identity for $0 \leqq t < \infty$ (the traditional representation of the Stieltjes transform results from the integral (15) after the substitution $(t,s) \rightarrow (t^{\frac{1}{2}}, s^{\frac{1}{2}})$ and a renaming of f).

First, Fourier inversion of (1) gives

$$(16) \qquad \tfrac{1}{2}\pi e^{-t^a} = \int_0^\infty F(x) \cos tx\, dx \qquad (0 \leqq t < \infty).$$

Next, if (12) is inserted in (16), the order of the integrations can be interchanged (by uniform convergence, if $t > 0$). Since

$$(17) \qquad \int_0^\infty e^{-tx} \cos sx\, dx = t/(t^2+s^2)$$

(if $t > 0$), it follows that (15) will be satisfied if (and, in view of the uniqueness theorem of Stieltjes' transform, only if) the function $h(s)$ is chosen as follows:

$$(18) \qquad \tfrac{1}{2}\pi h(s) = s \sin(\mu s^a) \exp(-\lambda s^a).$$

From the *Unterfunktion*, (18), of the Stieltjes analysis, (15), of the function e^{-t^a} the *Unterfunktion*, say $f(s)$ of its Laplace analysis

$$(19) \qquad e^{-t^a} = \int_0^\infty e^{-ts} f(s)\, ds,$$

where $0 \leqq t < \infty$, can be obtained, with the following result: If the constants λ, μ are defined by (13), and if (11) is assumed, then the f for which (19) holds for $0 \leqq t < \infty$ as an identity is given by

$$(20) \qquad \tfrac{1}{2}\pi f(s) = \int_0^\infty e^{-\lambda x^a} \sin \mu x^a \sin sx\, dx,$$

It may be mentioned that, while (11), (13) and (20) hardly indicate that

$$(21) \qquad f(s) \geqq 0 \text{ for } 0 \leqq s < \infty,$$

(21) happens to be true. In fact, successive differentiations of the function e^{-t^a} readily shows that, if $0 < t < \infty$ and $0 < a < 1$, its n-th derivative is positive or negative according as n is even or odd. Hence the Hausdorff-Bernstein theorem on completely monotone functions, when combined with the uniqueness theorems of Laplace's transform, implies that (21) is a consequence of (19).

9. The deductions of (20) from (18) has hardly anything to do with the explicit form of the function e^{-t^a}; it depends merely on the following formal rule: If

$$(22) \qquad \int_0^\infty (t^2+s^2)^{-1}h(s)\,ds = \int_0^\infty e^{-ts}f(s)\,ds$$

is an identity for $0 \leqq t < \infty$, then (under appropriate conditions, conditions which are amply satisfied in the case of e^{-t^a}) the connection between f and h is given by Dirichlet's transform,

$$(23) \qquad f(s) = \int_0^\infty h(x) \sin sx / x\, dx. \qquad (0 < s < \infty).$$

Hence (20) follows from (18). (Incidentally, if f, rather than h, is known in (22), then, subject to the legitimacy of the Fourier inversion which is involved,

$$(24) \qquad \tfrac{1}{2}\pi h(s)/s = \int_0^\infty f(x) \sin sx\, dx, \qquad (0 < s < \infty);$$

in fact, (23) is the Fourier sine transform of $h(s)/s$, and so (24) follows by Fourier reciprocity).

I cannot decide whether the solution (23) of (22) is generally known; in view of [2], something like (22)-(23) must have been familiar to Ramanujan. In any case, it would be worthwhile to exploit this rule systematically, *by completing the tables of known Laplace transforms (for a given* Unterfunktion *f) by those items for which the* Unterfunktion *h of the Stieltjes transform is already recorded in the tables of the latter transform.* The actual validity of the Dirichlet rule (23) in the appropriate Hilbert space (that is, the specification of such function spaces) can be justified along the lines of the Plancherel technique of Paley and Wiener [9] (*passim*; e.g., pp. 42-43). In such cases as e^{-t^a}, neither locally nor in Hilbert's space is there an issue, and (20) could be deduced from (15) by an appeal to

Fubini's theorem. But it is somewhat more instructive to verify (22) under the drastic hypothesis that it is legitimate "to compare like powers."

To this end, note that, if $(t^2+s^2)^{-1}$ is expanded according to powers of $1/t$, then, after a term-by-term integration (if it is allowed), the integral on the left of (22) appears in the form

$$\sum_{n=0}^{\infty}(-1)^n\mu_{2n}(h)/t^{2(n+1)}, \text{ where } \mu_n(h)=\int_0^\infty s^n h(s)\,ds.$$

On the other hand, successive partial integrations lead to the following expansions of the integral on the right of (22):

$$\sum_{n=0}^{\infty} f^{(n)}(0)/t^{n+1}, \text{ where } f^{(n)}=d^nf/ds^n$$

(if $f(s)$ is of class C^∞). Hence, comparison of the respective powers of $1/t$ (if it is allowed) shows that

$$f^{(2n)}(0)=0 \text{ and } f^{(2n+1)}(0)=(-1)^n\int_0^\infty s^{2n}h(s)\,ds,$$

where $n=0,1,\cdots$. Thus, if $f(s)=\sum_{n=0}^{\infty} f^{(n)}(0)s^n/n!$, it follows that

$$f(s)=\sum_{n=0}^{\infty}(-1)^n s^{2n+1}/(2n+1)!\int_0^\infty x^{2n}h(x)\,dx,$$

which is (23) if $\Sigma\int=\int\Sigma$.

10. The suggestion italicized above, concerning the explicit formulation (23) (or, in the reverse direction, (24)) of the hypothesis (22), is illustrated by the transition from (15) and (18) to (19) and (20). In view of (17), this can be re-stated so as to replace the Stieltjes transform by a Fourier *cosine* transform (whereas (23) or (24) is a Fourier *sine* transform); in fact, the e^{-x^a} in (1) and (16) can be replaced by any function which is "small and smooth." Actually, all of this can be formulated as a general principle, concerning the translation of one transform table into another transform table, as follows:

For $M=K,M,N,\cdots$, let $M=M(t,s)$ be the kernel of an integral transformation

$$\int_0^\infty M(t,s)g(s)\,ds \tag{25}$$

which sends "arbitrary" functions g on $0 < s < \infty$ into "arbitrary" functions $g^* = Mg$ on $0 < t < \infty$ and which has the property that (in terms of an appropriate definition of identity) the *M-Unterfunktion*, $M^{-1}g^*$, of g is unique. With reference to two such transformations, M and N, let there be known a $K = K_{MN}$ for which $M^{-1} = KN^{-1}$ holds (in view of the formulation (23) of (22), this is the case if $M = L$, $N = S$ and $K = D$, where L, S and D denote the transforms of Laplace, Stieltjes and Dirichlet respectively, with (t, s) replaced by (t^2, s^2) in the traditional representation of S). Then, if the *N-Unterfunktion* of a *given* function of t happens to be known explicitly, the *N-Unterfunktion* of the *same* function of t follows by an integration, that assigned by K.

Instances and general methods are contained in [2] and [9].*

Let the function (25) of t on $(0, \infty)$ be denoted by $S_t f$, $L_t f$, $C_t f$ in the respective cases $M = S, L, C$, where

$$(26_S) \quad S = (t^2 + s^2)^{-1}; \qquad (26_L) \quad L = e^{-ts}; \qquad (26_C) \quad C = \cos ts.$$

Note that D is the integral (vanishing at the origin) of C; cf. (23).

By Fourier's inversion, $\frac{1}{2}\pi C^{-1} = C$ (formally; so that, in particular, (17) is equivalent to

$$(17\text{ bis}) \qquad \int_0^\infty (t^2 + x^2)^{-1} \cos sx \, dx = \tfrac{1}{2}\pi e^{-ts}/t,$$

where $t \neq 0$.) It follows that if

$$(27) \qquad \int_0^\infty |f(s)| \, ds < \infty,$$

then, for $0 < t < \infty$,

$$(28) \qquad (LC)_t f = S_t f_1, \text{ where } f_1(s) = f(s)s,$$

and that (28) has the dual

$$(28\text{ bis}) \qquad (SC)_t f = \tfrac{1}{2}\pi L_t f_{-1}, \text{ where } f_{-1}(s) = f(s)/s,$$

* For specific pairs M, N of certain type, conditions under which this obvious remark is legitimate (for a given f) can be read off, for instance, from such considerations as were collected by Hirschman and Widder in their book [5] (in which, incidentally, it is again overlooked, as in the other publications of these authors and in those of I. J. Schoenberg which they quote, that the Fourier representation of the reciprocal Weierstrass products was studied already in [12], pp. 51-54 and 61-64, where, in the even case, these functions are introduced precisely in the same way as then followed by I. J. Schoenberg, cf. [5]).

if, in addition to (27),

$$\int_{+0} |f(s)/s|\, ds < \infty. \tag{27 bis}$$

In fact, the content of (28) has already been used. But in view of (17) and its Fourier inversion (17 bis), both (28) and (28 bis) follow from Fubini's theorem. As mentioned above, there is a corresponding deduction of (23) from (22). It is also clear that, under the respective Fubini assumptions, (27) or (27) *and* (27 bis), of (28) and (28 bis),

$$LC = CL \text{ in (28), and } SC = CS \text{ in (28 bis).} \tag{29}$$

The C-duality, (28)-(28 bis) and (29), of L and S has an analogue, but becomes involutory, if (26_C) is retained but (26_L) is replaced by the "normal" (Gaussian) kernel

$$N = e^{-t^2 s}. \tag{26_N}$$

What then corresponds to (26_S) is, in the main, (26_N) itself, since what corresponds to (17) *and* to (17 bis) is contained in the *single* relation

$$\int_0^\infty e^{-x^2} \cos yx\, dx = \tfrac{1}{2}\pi^{\frac{1}{2}} e^{-z^2}, \text{ where } z = \tfrac{1}{2}y. \tag{30}$$

While this parallelism is trivial, it turns out that, owing to the (N, N)-analogue of the (L, S)-rules (28)-(29), the explicit results (16)-(20), where (11) and (13) are assumed, settle a desideratum I formulated a long time ago, concerning the actual determination of the weight functions which, *for all values of* β *on the range* $0 < \beta < 2$, are the factors of normal stratification ("Gaussian analysis") of Cauchy's (symmetric) *stable distribution functions of index* β; cf. [4], Section 4, where further references are given (p. 769, footnote [20]).

11. In order to avoid a confusion of the index a, occurring in (11) and (13), with Cauchy's index β, let the functions F, G, h, f, occurring in (1)-(3) and (18)-(20), be now denoted by F_a, G_a, h_a, f_a, and let $\beta = \beta(a)$ be an abbreviation for

$$\beta = 2a; \text{ so that } 0 < \beta < 2, \tag{31}$$

by (11). Thus F_γ, G_γ are defined, and (1)-(3) and (16) remain valid, if $\gamma = a$ is replaced by $\gamma = \beta$, but (7), (9), (12), (15) and (19)-(20) then

become divergent (and, incidentally, the analogue of (21) becomes wrong). The notation $\lambda=\lambda(\gamma)$, $\mu=\mu(\gamma)$ will always be used only with reference to $\gamma=\alpha$ (so that the numbers (11) will remain positive).

Except for a trivial factor $(=1/\pi)$, Cauchy's (symmetric) "stable freqency functions" of index β is $F_\beta(t)$, which means, by (1), that

$$F_\beta(t)=\int_0^\infty e^{-s^\beta}\cos ts\,ds. \tag{32}$$

On the other hand, if the t in (19), where $0\leqq t<\infty$ and $0<\alpha<1$, is replaced by t^2, it follows from (31) that

$$e^{-t^\beta}=\int_0^\infty e^{-t^2 s}f_\alpha(s)\,ds. \tag{33}$$

It follows therefore from (32) (and from Fubini's theorem) that

$$F_\beta(t)=\int_0^\infty f_\alpha(s)\{\int_0^\infty e^{-x^2 s}\cos tx\,dx\}\,ds.$$

But here the { } is $\frac{1}{2}\pi^{\frac{1}{2}}\exp(-t^2/4s^2)$, as seen by changing the integration variable in (30). Accordingly

$$F_\beta(t)=\tfrac{1}{2}\pi^{\frac{1}{2}}\int_0^\infty \exp(-t^2/4s^2)f_\alpha(s)\,ds. \tag{34}$$

The relation (34) is the explicit form of the stratification, announced above. In fact, the weight function occurring in (34) is

$$f_\alpha(s)=2\pi^{-1}\int_0^\infty \sin sx\,e^{-\lambda x^\alpha}\sin\mu x^\alpha\,dx\geqq 0,\quad \text{cf. (11)}, \tag{35}$$

by (20) and (21). Incidentally, the inequality in (35) makes it clear that the function (34) of t^2 is not only positive but decreasing as well. This property of Cauchy's stable densities $F_\beta(t)$, where $0<\beta<2$, is significant from the point of view of the theory of probability (cf. [12], pp. 70-71); it was proved in [12], pp. 83-86, by arguments having the nature of an existence proof, instead of being explicit, as in (34)-(35).

Another "stratified" representation of $F_\beta(t)$ (which, however, fails to exhibit the essential information $f_\alpha\geqq 0$ in (34)) is the following:

$$F_\beta(t)=\pi^{-\frac{1}{2}}\int_0^\infty E(\tfrac{1}{2}t,x)\,\mathrm{Im}\exp(-\lambda+i\mu)x^\alpha\,dx, \tag{36}$$

where $\lambda=\lambda_\beta>0$, $\mu=\mu_\beta>0$ and $a=a_\beta<1$ are defined by (11) and (31), and $E(t,s)$ is the function

$$(37)\qquad E(t,x)=\operatorname{Im}\int_0^\infty \exp(ixs-t^2/s^2)\,ds,$$

which is independent of β (<2). In fact, if (35) is inserted in (34) and the order of integrations is interchanged (which can readily be justified), then (36) follows from (37).

12. If $F(t)$ is any *symmetric frequency curve* $(=F(-t)\geqq 0$, where $-\infty<t<\infty)$, it is called *unimodal* if $F(t)$ is a non-increasing function of $t>0$, and it is called *bell-shaped* if it is such that there exists a (not necessarily unique) positive t_0 having the following property: The curve $F=F(t)$ is concave or convex (toward the t-axis) according as $0\leqq t\leqq t_0$ or $t_0\leqq t<\infty$. Both of these classes of frequency curves F, the second of which is a subclass of the first, can be defined without the assumption $F(t)=F(-t)$ also. The unimodal character of (34), just concluded from the inequality in (35), can of course be concluded also if (34) is generalized to

$$(38)\qquad F(t)=\int_0^\infty \exp(-t^2x)\,d\phi(x),$$

$(x=1/4s^2)$, where $\phi(x)$ is any function satisfying

$$(39)\qquad \int_0^\infty d\phi(x)=\int_0^\infty |\,d\phi(x)|<\infty,\ \text{i.e., } d\phi(x)\geqq 0,\ \phi(\infty)-\phi(0)<\infty.$$

The Gaussian limiting case, $\beta=2$, of (34) results from (36) if $\phi(x)$ is chosen to be the step-function $\phi_{pa}(x)=p\operatorname{sgn}(x-a)$, where p and a are positive constants. But whereas (38) is bell-shaped if $\phi=\phi_{pa}$, it is possible to choose four positive constants $p, a;\ q, b$ in such a way that the case $\phi=\phi_{pa}+\phi_{qb}$ of (38), the superposition of two symmetric normal frequency curves, fails to become bell-shaped; cf. [11]. Thus, whereas one might hope that (34) is bell-shaped not only in the limiting case $\beta=2$ but for every positive $\beta<2$, this, if true, cannot be concluded from nothing more than the inequality in (35).

If $\beta=1$, then Cauchy's frequency function (34) is known to reduce to the elementary function $(1+t^2)^{-1}$ (except for constant factors) and is therefore bell-shaped. Since this means that $d^2F_\beta(t)/dt^2$ has a (unique)

positive zero $t_0 = t_0(\beta)$ if $\beta = 1$, it can be concluded from (34) and from the explicit form (35) of $d\phi(x)/dx$ that such a $t_0(\beta) > 0$ exists not only for $\beta = 1$ but for every β contained in the range $1 - \epsilon < \beta < 1 + \epsilon$, where $\epsilon > 0$ is a sufficiently small explicit constant. But even this continuity argument can be justified only by a certain amount of explicit labor, and the desideratum (namely, that the entire range, $0 < \beta < 2$, of Paul Lévy can be included, so that ϵ can be chosen to be 1) cannot apparently be obtained along these lines.*

13. Let α in (1) now be any positive index (not restricted by $0 < \alpha < 1$ or $0 < \alpha < 2$). Then (4) is still valid, as is (3) with (2). If $\alpha > 1$, then the series (7) is divergent at every $t > 0$. But it can be seen from the proof of (7) that (7) holds, as $t \to +0$, as an asymptotic development,

$$(41) \qquad G(t) \sim \sum_{k=0}^{\infty} c_{k-1} t^{k-1}/\Gamma(k), \qquad c_k = (-1)^{k-1} \alpha \Gamma(\alpha k) \sin(\tfrac{1}{2}\pi\alpha k).$$

In view of (2), this is an elaboration of (4), where $t \to \infty$. If α is an even integer, $\alpha = 2n$, then all coefficients of (41) become 0, hence (41) means that, for every fixed $\epsilon > 0$,

$$(42) \qquad G(t) = O(t^{1/\epsilon}) \text{ as } t \to 0 \qquad (\alpha = 2n)$$

and so, in view of (2),

$$(43) \qquad F(t) = O(t^{-1/\epsilon}) \text{ as } t \to \infty \qquad (\alpha = 2n)$$

(actually, the by-product (42) of (41) can be obtained directly; in fact, it is readily seen from (3) that all derivatives of the function $G(t)$, which clearly is analytic for $0 < t < \infty$, tend to 0 as $t \to 0$, if $t > 0$).

If $2n = 2$, then (43) is obvious, since (1) reduces to

$$F(t) = \tfrac{1}{2}\pi^{\frac{1}{2}} \exp(-t^2/4) \text{ if } \alpha = 2.$$

---

* It is known that if $\Phi(x) = \Phi(-x)$, where $-\infty < x < \infty$, denotes the Fourier transform of Riemann's $\Xi(t) = \xi(s) = \Xi(-t)$, where $s = \frac{1}{2} + it$, then not only $\Phi(x) > 0$ holds (Jensen, Hurwitz) but also $d\Phi(x)/dx < 0$, where $0 < x < \infty$ (see my note in the Journ. London Math. Soc., vol. 10 (1935), pp. 82-83). Is it also true that $\Phi(x)$ is bell-shaped?

Incidentally, it is clear from the two known properties of $\Phi(x)$ that the first of the functions

$$(40) \qquad \int_0^\infty \Phi(x) \sin tx \, dx, \qquad \int_0^\infty \Phi(x) \cos tx \, dx$$

is positive for $0 < t < \infty$, whereas the second changes sign an infinity of times, since $\Xi(t)$ does.

One will therefore expect that, if

$$(44) \qquad a = 2n, \text{ where } 2n = 4, 6, 8, \cdots,$$

then $F(t)$ will be much smaller than the order supplied by (43). But after the exclusion of the trivial case $2n = 2$, the function $F(t) = F_{2n}(t)$ changes sign an infinity of times as $t \to \infty$ (in this regard, cf. the known results of Pólya, referred to at the end of Section 2), and what one will conjecture for (44) is the existence of an explicit asymptotic formula

$$(45) \qquad F(t) \sim f(t) \cos g(t) \text{ as } t \to \infty,$$

where $f(t)$ and $g(t)$ are elementary functions which satisfy $-\log\log f(t) \to \text{Const.} > 0$ and $\log g(t) \to \text{const.} > 0$, where $t \to \infty$ (and where Const. and const. depend on the index (44)). In a letter to Professor Pólya, I raised this question a long time ago, mentioning the analogy between the conjectural formula (45) and the known formulae for the "(real) definite integral" solutions of the confluent forms of hypergeometric equation (Bessel-Poisson; Laguerre-Perron); the case $a = 2n$ of being a "Laplace" solution of a simple differential equation, of order $n-1$ which is homogeneous and linear:

$$(46) \quad d^{n-1}F(t)/dt^{n-1} + atF(t) = 0, \text{ where } a = a_n = \text{const.} = (-1)^{n-1} | a_n |.$$

In another context (not that of Cauchy's "stable densities"), the identity (46) for the case $a = 2n$ of (1), readily verified by successive partial integrations, is known since Jacobi's time (at least); cf., e.g., p. 152 of E. G. C. Poole's *Linear Differential Equations* (Oxford, 1936).

Professor Pólya was good enough to determine the explicit form of (45) and to communicate it to me in a letter dated November 18, 1944, the content of which is only now published on his request. The formula is of an impressive length; I shall write it in terms of notations corresponding to those used in (12) (where (11), rather (44), is the assumption).

In terms of a fixed n (> 1), put $\nu = (n-1)/(2n-1)$, $\kappa = 2n/(2n-1)$ (hence $\nu > 0$) and

$$\lambda = -\rho \cos(\tfrac{1}{2}\pi\kappa), \quad \mu = -\rho \sin(\tfrac{1}{2}\pi\kappa), \text{ where } \rho = (2n-1)/(2n)$$

(so that $\lambda > 0$, since, in view of (44), the angle $\frac{1}{2}\pi\kappa = \pi n/(2n-1)$ is between $\frac{1}{2}\pi$ and π). Then, if $a = 2n$ and $t \to \infty$ in (1),

$$F(t) \sim Ct^{-\nu} \exp(-\lambda t^{\kappa}) \cos(\mu t^{\kappa} + \tfrac{1}{2}\pi\nu),$$

where $C = (8\pi/c)^{\frac{1}{2}}$, $c = m(1+m)^{1/m}$, $m = 2n-1$.

Pólya added in his letter that he could foresee this structure of $f(t)$ and $g(t)$ in (45) from the order of magnitude of purely imaginary values, which can be ascertained from the Maclaurin series of (1) and from (46); and that he obtained the final result by writing the case (44) of (1) in the form

$$\int_{-\infty}^{\infty} \exp\{itx - x^{2n}\}dx \equiv t^{1/m} \int_{-\infty}^{\infty} \exp\{(ix - x^{2n})t^{2n/m}\}dx, \tag{14 bis}$$

where $m = 2n - 1$ and $0 < t < \infty$, and applying the method of steepest descent to the last integral.

APPENDIX I.*

Reduction mod 1.

Let $0 < \alpha < \infty$, $0 < q < 1$, $-\infty < x < \infty$, and put

$$\theta_\alpha(x; q) = 1 + 2\sum_{m=1}^{\infty} q^{m^\alpha} \cos 2\pi mx; \tag{1}$$

so that $0 \leqq x < 1$ without loss of generality. This is Poisson's kernel if $\alpha = 1$, and an elliptic theta function if $\alpha = 2$.

If $\alpha = 2$, then the function

$$F_\alpha(t) = \int_0^{\infty} \exp(-s^\alpha) \cos ts\, ds = F_\alpha(-t) \qquad (0 \leqq t < \infty) \tag{2}$$

differs from its Fourier reciprocal (given by

$$\pi \exp(-|t|^\alpha) = \int_0^{\infty} F_\alpha(s) \cos ts\, ds \tag{3}$$

for any $\alpha > 0$) only in two constant factors. Hence, when the functional equation (the "linear transformation") of the case $\alpha = 2$ of (1) is proved by means of Poisson's summation formula

$$\sum_{k=-\infty}^{\infty} \phi(x + k) = \sum_{m=-\infty}^{\infty} e^{2\pi imx} \int_{-\infty}^{\infty} \phi(y) e^{-2\pi imy}\, dy, \tag{4}$$

it is immaterial whether $\phi(rt)$, where r is any positive constant, is chosen

* Added July 21, 1956.

to be the function (2) or the function (3). On the other hand, the two choices lead to two different identities in every non-involutory case ($\alpha \neq 2$).

The first choice of ϕ leads to the following representation of (1) (for any $\alpha > 0$):

$$2\pi\theta_\alpha(x;q) = p \sum_{k=-\infty}^{\infty} F_\alpha(px+pk), \tag{5}$$

where $p = p(q) = p_\alpha(q)$ is defined by

$$(-\log q)^{1/\alpha} = 2\pi/p \tag{6}$$

(so that the range $0 < q < 1$ is mapped on the half-line $0 < p < \infty$). In fact, if $\phi(x) = F_\alpha(px)$, where $-\infty < x < \infty$ and $p = \text{const.} > 0$, then, since (2) is equivalent to

$$2\pi p^{-1} \exp(-|x|^\alpha/p^\alpha) = \int_{-\infty}^{\infty} F_\alpha(py) e^{-ixy}\, dy,$$

the series on the right of (4) becomes $2\pi p^{-1} \sum e^{2\pi imx} \exp(-|2\pi m|^\alpha/p^\alpha)$, which is $2\pi p^{-1}\theta_\alpha(x;q)$, by (1) and (6). Hence (5) follows by inserting $\phi(x) = F_\alpha(px)$ on the left of (4) also.

If α is of the form $2n$, then (2) is majorized by $\exp(-a|t|^b)$, where $a = a(\alpha) > 0$ and $b = b(\alpha) > 0$ (cf. Section 13), and so it is clear that

$$\sum_{k=-\infty}^{\infty} F_\alpha(px+pk) \sim F_\alpha(px) \text{ as } p \to \infty \tag{7}$$

holds at every fixed $x > 0$, and uniformly for $\epsilon < x < 1/\epsilon$ if $\epsilon > 0$ is fixed. It follows therefore from (5) and (6) that, for $0 < x < 1$, and uniformly for $\epsilon < x < 1 - \epsilon$ (< 1),

$$\theta_\alpha(x;q) \sim |\log q|^{-1/\alpha} F_\alpha(2\pi |\log q|^{-1/\alpha} x) \text{ as } q \to 1. \tag{8}$$

In the classical case, where $\alpha = 2$, the asymptotic formula (8) becomes dull, since $F_2(t) = \pi^{-\frac{1}{2}} \exp(-t^2/4)$ is free of fluctuations. But if $\alpha = 2n > 2$, then, by Pólya's formula,

$$F_\alpha(t) \sim Ct^{-\nu} \exp(-\lambda t^\kappa) \cos(\mu t^\kappa + \tfrac{1}{2}\pi\nu) \text{ as } t \to \infty \tag{9}$$

holds for the (even) function (2), and C and the four Greek indices occurring on the right of (9) are certain (non-vanishing) functions of n (> 1) alone (cf. Section 13), and (8) and (9) (where $\lambda > 0$ and $\nu > 0$) show that, as $q \to 1$, the function (1) of q is wobbly at every x distinct from 0 (mod 1), and uniformly for $\epsilon < x < 1 - \epsilon$ (< 1).

If α is not an even integer, then the asymptotic behavior ($q \to 1$) of

(1) is as dull as it is in the case $a=2$. But (8) is then wrong, since the terms of the series (5) are damped (for increasing $|k|$) with sufficient rapidity only if $a=2n$ (where $a=1$ is allowed). In fact, if $0<a<\infty$ but $a\neq 2n$, then $F_a(t)=F_a(-t)$ is asymptotic to $A/|t|^{1+a}$ as $t\to\infty$, where $A=A(a)$ is a non-vanishing constant (cf. formula (4) of Section 2). Hence, if $p\to\infty$, then, for fixed $x>0$ (and uniformly for $\epsilon<x<1/\epsilon$), the function on the left of (7) is asymptotic to $A\psi_a(x)/p^{1+a}$, where

$$(10^*)\qquad \psi_a(x)=\sum_{k=-\infty}^{\infty}|x+k|^{-1-a}=\psi_a(x+1)\qquad (a>0).$$

It follows therefore from (5) and (6) that, uniformly for $\epsilon<x<1-\epsilon\,(<1)$,

$$(10)\qquad \theta_a(x;q)\sim B\psi_a(x)\log q \text{ as } q\to 1,$$

where $B=-A/(2\pi)^{1+a}$, $A=\Gamma(1+a)\sin(\tfrac{1}{2}\pi a)$.

If $x=0$, then, although (5) is valid, (8) is false, and (10) is meaningless (because (10*) includes $k=0$). But if $x=0$ and $q\to 1$, then there is no issue, since (1) shows that, if $r=-\log q$ (hence $0<r\to 0$), then $\tfrac{1}{2}\theta_a(0;q)-\tfrac{1}{2}$ is

$$\sum_{m=1}^{\infty}\exp(-rm^a)\sim I_a(r),\text{ where } I_a(r)=\int_0^{\infty}\exp(-ru^a)\,du,$$

and since $I_a(r)=I_a(1)/r^{1/a}$, where $I_a(1)=\Gamma(1+1/a)$.

Consider finally (1) as a function of x at a fixed q ($\neq 0$; so that the case $\theta_a(q;x)=$ const. will now be excluded). Clearly, $\theta_a(q;x)$ is of class C^∞ on $0\leqq x<1 \pmod 1$ whenever $a>0$, is analytic at every real x if $a=1$, and is a transcendental entire function of the complex variable x if $a>1$. If $a<1$, then (1) cannot be analytic for all real x, since the coefficient of the m-th cosine fails to be majorized by the m-th term of a convergent geometrical progression. But (1) is analytic on the interval $0<x<1$ (with a singularity of the C^∞-type, like $\exp(-x^{-2})$, at $x=0$) if $a<1$. This can be seen as follows:

If the $1+2\Sigma$ of (1) is simplified to the sum Σ itself, then what results is the real part of $P(z)=\Sigma q^{m^a}z^m$ on $|z|=1$ ($\arg z=2\pi x$), where $m=1, 2,\cdots$. Hence the assertion is that this power series $P(z)$ (which, since $a<1$, diverges if $|z|>1$) converges for $|z|<1$ to a function which remains regular at every point $z\neq 1$ of $|z|=1$. Actually, the analytic continuation of $P(z)$ exists (as a single-valued regular function) on the domain which results if the half-line $1\leqq z<\infty$ is removed from the z-plane. In fact, the coefficient of z^m in $P(z)$ is of the form $f(m)$, where $f(w)$ is an

entire function which, since $|q|<1$ and $0<\alpha<1$, is $O(\exp|w|^{\epsilon})$ as $|w|\to\infty$, if $\epsilon>0$ is fixed. Hence the assertion follows from a known general theorem (cf. [7], p. 109).

In order to exhibit the precise nature of the singularity of (1) at $x=0$ (if $0<\alpha<1$ and $0<q<1$), let the expansion (7) of Section 5, an expansion valid for all t, be inserted in formula (2) of Section 2. This shows that

$$F_\alpha(t)=\sum_{j=1}^{\infty} b_j/t^{1+\alpha j} \text{ for } 0<t<\infty, \tag{11}$$

where the coefficients $b_j=b_j(\alpha)$ are given by

$$b_j=c_{j-1}/\Gamma(j), \qquad c_{j-1}=(-1)^{j-1}\alpha\Gamma(\alpha j)\sin(\tfrac{1}{2}\pi\alpha). \tag{11 bis}$$

On the other hand, since (6) is positive and (2) is even, (5) can be written in the form

$$2\pi\theta_\alpha(x;q)=p\sum_{k=0}^{\infty}F_\alpha(pk+px)+p\sum_{k=1}^{\infty}F_\alpha(pk-px),$$

if $x\geqq 0$. Hence, if the F_α of the latter two series are substituted from (11) and the summation order is interchanged (which is readily justified, since the coefficients (11 bis) decrease rapidly enough, and since $F_\alpha(t)=O(t^{-1-\alpha})$ as $t\to\infty$), it is seen from (6) that

$$2\pi\theta_\alpha(x;q)=\sum_{j=1}^{\infty} b_j\chi_\alpha^{\,j}(x)/(-2\pi)^{\alpha j} \tag{12}$$

holds for $0<x<1$ and for every $\alpha<1$, where the functions $\chi_\alpha^{\,1}(x),\chi_\alpha^{\,2}(x),\cdots$ denote the following counterparts of the functions (10*):

$$\chi_\alpha^{\,j}(x)=\sum_{k=0}^{\infty}(k+x)^{-1-\alpha j}+\sum_{k=1}^{\infty}(k-x)^{-1-\alpha j}. \tag{12*}$$

The representation (12) of (1), a representation which in view of (12*) is (in an appropriate sense) of the type of an Eisenstein series (for a modular function), puts into evidence the nature of the singularity acquired by (1) when x tends to 0 (mod 1) if $\alpha<1$. This "decomposition into singularities" was obtained in [19] by the above method but with algebraic errors, corrected in (12) and (12*) above (the errors were introduced in [19] by an erroneous copying (from [13]) of the exponents in (11) above, that is, in the [correct] result of [13]).

Another application of (5) results if use is made of formulae (31)-(35) of Section 11. According to those formulae, which are valid for $F_\alpha(t)$ if $\alpha<2$ (rather than only if $\alpha<1$),

$$\pi^{\frac{1}{2}}F_a(t)=\int_0^\infty \exp\left(-t^2/4s^2\right)\phi_a(s)\,ds, \tag{13}$$

if

$$\phi_a(s)=\int_0^\infty \sin(sy)\exp(-\lambda y^{\frac{1}{2}a})\sin(\mu y^{\frac{1}{2}a})\,dy, \tag{14}$$

where, corresponding to the definitions (11) in Section 7, $\lambda=\cos(a\pi/4)$, $\mu=\sin(a\pi/4)$. It is clear from (5) and (13) that

$$2\pi^{\frac{1}{2}}\theta_a(x;q)=\int_0^\infty \phi_a(s)\Theta(x;s)\,ds, \qquad (0<a<2), \tag{15}$$

where, if $p=p_a(q)$ is defined by (6),

$$\Theta(x;s)=p\sum_{k=-\infty}^{\infty}\exp\{-(px+pk)^2/4s^2\}. \tag{15*}$$

But the case $a=2$ of (5) shows that (15*) is of the form $\theta_2(x;r)$, where θ_2 is an elliptic theta and $r=r(s,p)=r_a(s;q)$. Since the function (14) is positive for $0<s<\infty$ (cf. (20)-(21), Section 8), it follows that (15) is the analogue for "angular distributions" of that "Gaussian stratification" the explicit form of which was obtained in Section 11.

It was Zernike's discussion of angular statistics ([20], pp. 477-478) which led me in [16] to the interpretation of the function (1) as the "densities of stable, symmetric, angular distributions." Subsequently, these considerations were rediscovered, and further developed (also in the unsymmetric case) by Lévy [15]. The proof of a certain property ("symmetrical convexity" mod 1) of the case $a=2$ of (1) is incomplete in [15], p. 36; a simple proof of that property was given in [17], and a refinement ("bell-shaped graph" mod 1; cf. Section 12 above) of that property in [18]. The existence of (15), without the explicit form, (14), of the weight function of (15), was shown in [17], where the weaker one of the two properties, just mentioned, was extended from $a=2$ to $0<a<2$ (that extension is contained in the fact that the weight function (15) is positive for $0<s<\infty$).

It should finally be mentioned that there is a fundamental difference between the "linear" and the "angular" case. If the unit of length is suitably chosen, then the stable distributions having an even density are given by (2), where, however, the index $a>0$ cannot be chosen arbitrarily (as proposed by Cauchy), since the density does not become negative if and only if $a\leqq 2$ (Lévy). In contrast, it is clear that there belongs to *every* a,

where $0 < a < \infty$, a sufficiently small $q^* = q^*_a > 0$ in such a way that the function (1) (in which the value of the parameter q has the role of a standard deviation) is positive for all x whenever $0 \leqq q < q^*$.

APPENDIX II.

On certain unimodal distributions.

In the nomenclature of [12], the result of [12], p. 78, refined to an explicit representation by Section 11 above, was as follows:

(†) *Every symmetric stable distribution (Cauchy-Lévy) is a convex distribution.*

The stable distributions are the simplest instances of the so-called L-distributions (cf. below). Hence (†) is contained in the following theorem (*), the proof of which is the first purpose of this appendix:

(*) *Every symetric L-distribution (Khintchine-Lévy) is a convex distribution.*

In their monograph [22], Kolmogoroff and Gnedenko formulate, and attempt to prove, a theoerm which would contain (*). But as observed by K. L. Chung, the proof is based on an erroneous lemma. Cf. the footnote at the end of Appendix III below. Correspondingly, the following proof of (*) will have to be based on a more involved adaptation of the proof of (†) in [12].

Let $F_\phi(t)$, where $-\infty < t < \infty$, denote the Fourier-Stieltjes transform,

$$F_\phi(t) = \int_{-\infty}^{\infty} e^{itx}\, d\phi(x), \tag{1}$$

of a distribution $\phi = \phi(x)$, that is, of a function defined for $-\infty < x < \infty$ in such a way that

$$\phi(-\infty) = 0, \quad \phi(\infty) = 1 \text{ and } d\phi(x) \geqq 0. \tag{2}$$

If $\phi(x)$, when normalized by $2\phi(x) = \phi(x+0) + \phi(x-0)$, is identical with the distribution $1 - \phi(-x)$, then $\phi(x)$ is called symmetric. By a convex distribution is meant a symmetric unimodal distribution, that is, a distribution which for $0 < x < \infty$ and $-\infty < x < 0$ (but necessarily at $x = 0$) has a (not necesarily continuous) density which is a monotone function of $|x|$ $(\neq 0)$.

The proof of (*) will be reduced to the following lemma: If $S(t)$ denotes the (even, entire) function

$$S(t) = \int_0^t (\sin u)^2/u\, du, \tag{3}$$

and if q is any positive constant, then the Fourier inversion of the case

$$F_\phi(t) = e^{-qS(t)} \qquad (-\infty < t < \infty) \tag{4}$$

of (1) defines a unimodal distribution $\phi = \phi_q = \phi_q(x)$. Since (4) is even, hence $\phi(x) = 1 - \phi(-x)$, this means that every $\phi = \phi_q$ is a convex distribution.

It is clear that if $\phi(x)$ is any unimodal distribution, and if p is any positive constant, then $\psi(x) = \phi(px)$ is a unimodal distribution. It is also clear that if $\phi = \phi(x)$ and $\phi_1 = \phi_1(x), \phi_2 = \phi_2(x), \cdots$ are distributions satisfying $\phi_n \to \phi$ as $n \to \infty$, then ϕ is unimodal whenever every ϕ_n is. Finally, a theorem of Hardy and Littlewood concerning "rearrangements" is substantially equivalent to the following fact (cf. [12], p. 47 and pp. 77-78): If ϕ and ψ are two *symmetric* unimodal distributions, then their convolution $\phi * \psi$ is a (symmetric) unimodal distribution.

In [12], the proof of (†) was based on the preceding three facts. The following proof of the generalization (*) of (†) will be an adaptation of the same proof (the first two of the three facts are needed, of course, only in the particular case of symmetry).

As will be shown in Appendix III below, Lévy's enumeration of all L-distributions (in terms of their Fourier-Stieltjes transforms; cf. [6], pp. 192-193, or [22], pp. 145-151) implies that a distribution $\phi(x)$ is a symmetric L-distribution if and only if there exists a sequence of distributions $\phi_1(x), \phi_2(x), \cdots$ satisfying $\phi_n \to \phi$ as $n \to \infty$, where every $\phi_n(x)$ is a convolution of the form

$$\phi_n(x) = \gamma_h(x) * \phi_{q_1}(p_1 x) * \cdots * \phi_{q_n}(p_n x), \tag{5}$$

in which $\gamma_h(x)$ denotes the symmetric normal distribution of a certain standard deviation $h = h_\phi$ ($\geqq 0$), the n numbers $p_i = p_i(n)$ are positive, the n numbers $q_i = q_i(n)$ are non-negative and, for every $q = q_i$, the distribution $\phi = \phi_q = \phi_q(x)$ is defined by (4) and (3), that is, by

$$\log F_\phi(t) = -q \int_0^t (\sin^2 u)/u\, du; \qquad -\infty < t < \infty. \tag{6}$$

[In view of (1), this implies that $\phi(x) = \frac{1}{2} + \frac{1}{2}\operatorname{sgn} x$ if $q = 0$; in what follows, this trivial distribution, which is certainly symmetric and unimodal, will be excluded, that is, q will be assumed to be positive.]

It now follows from the three facts, mentioned above, that the unimodal character of every symmetric L-distribution will be proved if (and only if) it is shown that (6) defines a unimodal distribution ϕ for every fixed $q > 0$. But it is clear from (6), where $F_\phi = F_{\phi_q}$, and from the product rule of the transforms (1) of convolutions, that $\phi_{q+r} = \phi_q * \phi_r$ whenever $q > 0$ and $r > 0$. Hence, the third of the three facts shows that, instead of proving the unimodal character of ϕ_q for every q on the range $0 < q < \infty$, it is sufficient to prove the same for every q on the range $0 < q < q_0$, where q_0 (> 0) can be chosen arbitrarily small.

Accordingly, it can be assumed that $0 < q < 1$. Then $\alpha = q/4$ and $\beta = 1 - q/2$ are positive numbers satisfying $\alpha + \beta + \alpha = 1$. Hence, if $\mu = \mu(x)$, where $\mu(-\infty) = 0$ and $\mu(\infty) = 1$, denotes the step-function having the jumps α and β at $x = \pm 2$ and $x = 0$ respectively, then μ is a distribution. Clearly, $F_\mu(t)$ is identical with the sum of $1 - \frac{1}{2}q$ and $\frac{1}{2}q\cos 2t$. Consequently,

$$F_\mu(t) = 1 - q\sin^2 t. \tag{7}$$

It is now easy to conclude that there belongs to $\phi = \phi_q$ a distribution $\lambda = \lambda_q$ for which

$$tF_\phi(t) = \int_0^t F_\lambda(s)\,ds \tag{8}$$

is an identity in t. In fact, if a prime denotes differentiation with respect to t, then $F_\phi'/F_\phi = -q(\sin t)^2/t$, by (6) and (5). In view of (7), this means that $(tF_\phi)' = F_\phi F_\mu$. Hence (8) is satisfied by the distribution λ which is the convolution of ϕ and μ.

Finally, since (6) is an even function of t, the distribution ϕ is symmetric, and so the assertion, according to which ϕ is a convex distribution, is equivalent to the statement that "the distribution ϕ is unimodal, with $x = 0$ as a mode." But a theorem of Khintchine states that this will be the case precisely if there exists some distribution λ satisfying (8) (in this regard, cf. chap III of Girault's thesis [21], where further references will also be found; concerning Khintchine's proof, cf. the presentation in [22], pp. 157-160, and K. L. Chung's comments on it in [22], pp. 251-253).

The unimodal character of the distribution $\phi = \phi_q$, defined by (6) and

(5), could be verified by an explicit determination (depending on a contour integration) of the Fourier inversion of $F_\phi(t)$. This would avoid an appeal to Khintchine's general criterion. The applicability of the latter is usually not straightforward, and it became straightforward above only because it was permissible to assume that q is sufficiently small.

The applicability of Khintchine's criterion (8) to the case (7) (for *small* q in $\phi=\mu$) seems to be one of the few instances in which the criterion happens to be amenable enough to lead to an affirmative result. In fact, the practical use of the criterion (which is a necesary and sufficient condition) lies perhaps in the opposite direction in most cases. This is illustrated by the proof of the following fact († bis):

(† bis) *If* $0<\alpha\leqq 2$, *then*

$$(1-\alpha\,|\,t\,|^\alpha)\exp(-|\,t\,|^\alpha) \qquad (-\infty<t<\infty) \tag{9}$$

is the Fourier-Stieltjes transform $F_\lambda(t)$ *of a (symmetric) distribution function* $\lambda=\lambda_\alpha(x)$.

In fact, (8) is equivalent to $(tF_\phi)'=F_\lambda$ or

$$F_\phi(t)+tF_\phi'(t)=F_\lambda(t). \tag{10}$$

On the other hand, the sum on the left of (10) becomes the product (9) if

$$F_\phi(t)=\exp(-|\,t\,|^\alpha). \tag{11}$$

Since the (non-trivial) symmetric stable distributions ϕ are characterized by (11), where $0<\alpha\leqq 2$, it follows that († bis) is equivalent to (†), the result quoted at the beginning of this Appendix. (Note that the rule, used at the end of the proof of (*), cannot be applied this time, since the first factor of the product (9), being unbounded as $t\to\infty$, is certainly not an F.)

A similar argument leads to the following curiosity:

If $\Xi(t)=\xi(\frac{1}{2}+it)$, where $\xi(s)$ *is Riemann's entire function, then there exists a* $p=p(x)\geqq 0$ *for which*

$$\Xi(t)+t\Xi'(t)=\int_0^\infty p(x)\cos tx\,dx, \tag{12}$$

where $\Xi'=d\Xi/dt$, *holds for* $-\infty<t<\infty$.

In view of the formulation (10) of Khintchine's criterion. this is equivalent to the result quoted in the footnote to Section 12. (The result of Jensen

and Hurwitz, referred to in that footnote, is the weaker statement that (12) holds for another $p \geqq 0$, say for $p^* = p^*(x)$, if the second term on the left of (12) is omitted.)

The assertion of († bis) (concerning Fourier *integrals*) is worth restating in terms of Fourier *series*. First, if $f(x)$ is the density of $\phi(x)$ in an absolutely continuous distribution function $\phi(x)$, then

$$\sum_{k=-\infty}^{\infty} f(x + 2\pi k) \tag{13}$$

represents (almost everywhere) a periodic function and, according to Poisson's summation formula, the Fourier series (L) of (13) is $(2\pi)^{-1}$ times

$$\sum_{n=-\infty}^{\infty} F_\phi(n) e^{-inx}. \tag{14}$$

But (14) is the Fourier series of a non-negative function,

$$\sum_{n=-\infty}^{\infty} (1 - a \mid rn \mid^c) \exp(- \mid rn \mid^a) e^{-inx},$$

and this Fourier series can be written in the form

$$1 + 2 \sum_{n=1}^{\infty} (1 + an^a \log q) q^{n^c} \cos nx \tag{15}$$

by placing $r^a = -\log q$. Since $0 < r < \infty$ means that $0 < q < 1$, it is seen that († bis) contains the following corollary:

If $0 < a \leqq 2$ *and* $0 < q < 1$, *then the function* (15) *is positive for all real* x. This fact is the more remarkable because the coefficient of $n^a (\to \infty)$ in (15) is a negative constant, $a \log q$ (an interpretation of the $\log q$ results if (15) is written in the form

$$\theta_a + a\partial\theta_a/\partial q, \qquad \theta_a = \theta_a(x; q), \tag{15 bis}$$

where θ_a is the function (1) of Appendix I if the x of (15) is replaced by $2\pi x$).

Actually, the italicized corollary of the formulation († bis) of (†) contains († bis) itself. This is seen by an application of the rule

$$\lim_{\epsilon \to 0} \sum_{m=1}^{\infty} \epsilon g(m\epsilon) = \int_0^{\infty} g(u)\, du$$

(Euler-Maclaurin), the provisos of which are amply satisfied in the present case (note that $q \to 1$ as $\epsilon \to 0$). On the other hand, if (15) is multiplied

by $\frac{1}{2}$ and if the argument of a complex variable z is denoted by x, then what results is the real part of

(16) $$\tfrac{1}{2}+\sum_{n=1}^{\infty}(1+\alpha n^{\alpha}\log q)\,q^{n^{\alpha}}z^{n}$$

on the circumference $|z|=1$. Consequently. († bis) is *equivalent* to the following theorem:

If $0<\alpha\leqq 2$ and $0<q\leqq 1$, then the real part of the power series (16) is positive in the circle $|z|<1$. In other words, the determinant conditions of the Carathéodory-Toeplitz criterion are satisfied by (16) for *every* positive $q<1$ if $0<\alpha\leqq 2$. An algebraic verification of the positivity of the determinants involved, leading to a direct proof of († bis), appears to be a hopeless task.*

APPENDIX III.

Symmetric *L*-distributions.

1. Let $\phi=\phi(x)$, where $-\infty<x<\infty$, be a distribution (in the sense that not only $d\phi(x)\geqq 0$ holds but also $\phi(-\infty)=0$ and $\phi(\infty)=1$), and let $F_{\phi}(t)$ denote its Fourier-Stieltjes transform,

(1) $$F_{\phi}(t)=\int_{-\infty}^{\infty}e^{itx}d\phi(x)\,;\quad -\infty<t<\infty.$$

Clearly, (1) goes over into its complex conjugate if $\phi(x)$ is replaced by the distribution $1-\phi(-x)$. If the latter is identical with $\phi(x)$, that is, if

(2) $$\phi(x)+\phi(-x)=1 \text{ for } -\infty<x<\infty$$

* Corresponding remarks hold if (16) is weakened to

(16 bis) $$\tfrac{1}{2}+\sum_{n=1}^{\infty}q^{n^{\alpha}}z^{n^{\alpha}}$$

("weakened" in the sense in which the second of the two functions (15 bis) is "weaker" than the first). In fact, a repetition of the preceding deduction shows that (i) the real part of the power series (16 bis) is positive in the circle $|z|<1$ whenever $0\leqq q\leqq 1$ and $0<\alpha\leqq 2$, and that (ii) the truth of (i) is equivalent to the statement Cauchy's symmetric stable distributions actually exist if $\alpha\leqq 2$; that is, to Lévy's result according to which the function (32) of Section 11 will not become negative if $\beta\leqq 2$. If the Carathéodory-Toeplitz criterion is applied to (16 bis), then (ii) leads to a purely algebraic formulation (but not of course to a proof) of Lévy's existence theorem; that is, of his result referred to under (ii).

(when $\phi(x)$ is normalized by $\phi(x)=\frac{1}{2}\{\phi(x+0)+\phi(x-0)\}$), then ϕ is called symmetric.

A fundamental result of Paul Lévy ([27]; cf. also [6], pp. 180-181, or [23], pp. 160-161) enumerates all function $F_\phi(t)$ for which $\phi(x)$ is an "infinitely divisible" distribution. An inspection of his general result shows that, in the particular case (2) (that is, if (1) is real-valued and/or even), the distribution is symmetric if and only if there exists on the closed half-line $0 \leqq u < \infty$ a function $\mu=\mu(u)$ satisfying

$$(3) \qquad \int_1^\infty u^{-2}\, d\mu(u) < \infty \text{ and } d\mu(u) \geqq 0 \text{ for } 0 \leqq u < \infty,$$

and having the property that (1) becomes representable in the form *

$$(4) \qquad F_\phi(t) = \exp\{-\int_0^\infty (\sin tu/u)^2 d\mu(u)\}.$$

Thus there is a one-to-one correspondence, $\phi=\phi^\mu$, between all symmetric, infinitely divisible distributions $\phi(x)$ and all functions $\mu(u)$ satisfying (3) if, for the sake of the one-to-one correspondence, μ is normalized by $\mu(u-0) = \mu(u)$ for $0 < u < \infty$. It is understood that the jump

$$(5) \qquad \mu(+0)-\mu(0)=c \geqq 0$$

need not be 0, and that (4) is meant to be

$$(6) \qquad F_\phi(t) = \exp\{-ct^2 - \int_{+0}^\infty (\sin tu/u)^2 d\mu(u)\}.$$

Lévy also succeeded in enumerating the distributions ϕ contained in the subclass of the infinitely divisible distributions which consist of L-distributions, the latter being the distributions defined by Khintchine in terms

---

* The distributions ϕ defined by (4) and (3), that is, by the particular case (2) of Lévy's result, are precisely the distributions rediscovered by J. von Neumann and I. J. Schoenberg [28] in their "metric geometry of Hilbert screws." This observation is made here because [28] as well as the subsequent publications consulted (cf., e.g., [26], pp. 434-438) on either subject (infinite divisibility, Hilbert screws) fail to mention the identity of the "metric" question with the particular case (2) of Lévy's result.

In this regard, the situation is particularly curious in the more recent text of Hille [26], who, *loc. cit.*, not only fails to mention Kolmogoroff's work on Hilbert space or the related corresponding considerations of [28] but it also ignores [25], even though it is precisely Lie's point of view of infinitesimal generators of Lévy's cyclic semi-groups which underlies [25].

of the addition of "asymptotically constant" random variables; cf. [6], pp. 192-193, or [22], pp. 145-151. An inspection of this result of Lévy shows that, in the particular case (2), a distribution $\phi(x)$ is an L-distribution if and only if it is a $\phi(x)=\phi^{\mu}(x)$ in which $\mu=\mu(u)$, besides satisfying (3), has the following properties: $\mu(u)$ is absolutely continuous for $0<u<\infty$ (while (5) can be positive) and, if $\mu'(u)$ denotes its density (almost everywhere), then (if $\mu'(u)$ is defined on the zero set in an appropriate way) the indefinite integral of $\mu'(u)/u^2$ is a monotone function of $\log u$ for $0<u<\infty$. It is readily seen that this is equivalent to

$$(7)\qquad d\mu(u)=\mu'(u)du \text{ and } d\{\mu'(u)/u\}\leqq 0, \text{ where } 0<u<\infty.$$

The object of the following considerations is an analysis of the class of symmetric L-distributions. Except for (i) in Section 2 (a result an extension of which to the unsymmetric case is not available; cf. the footnote at the end of Section 9 below), the symmetry assumption (2) is made only in order to simplify the formulae; the considerations will be such as to apply, *mutatis mutandis*, without the restriction (2) also. Incidentally, if $\phi(x)$ is any L-distribution, then $(1-\phi(-x)$ is an L-distribution and) the convolution of $\phi(x)$ and $1-\phi(-x)$ is a symmetric L-distribution; conversely, every L-distribution can be factorized in this manner. This is seen by inspecting the logarithms of the respective transforms (1).

2. In view of Theorem (v) below, a fact proved in Appendix II, implies the following Theorem (i):

(i) *Every symmetric L-distribution ϕ is unimodal.*

By this is meant that $\phi(x)$ is absolutely continuous for $0<x<\infty$ (hence, by (2), for $-\infty<x<0$), with a density which is a monotone function of $|x|$ (if $-\infty<x<\infty$ but $x\neq 0$; if $x=0$, then $\phi(x)$ can have a jump).

(ii) *A distribution ϕ is a symmetric L-distribution if and only if there belong to it a constant $c\geqq 0$ and, on the open half-line $0<u<\infty$, a function $\lambda=\lambda(u)$ satisfying*

$$(8)\qquad \lambda(u)\geqq 0,\ d\lambda(u)\leqq 0,\ \int_1^\infty \lambda(u)/u\,du<\infty,\ \int_{+0}^1 \lambda(u)du<\infty,$$

and leading to the following representation of the transform (1) *of* $\phi(x)=\phi_c^{\lambda}(u)$:

$$(9)\qquad F_\phi(t)=\exp\{-ct^2+\int_{+0}^\infty S(tu)d\lambda(u)\},$$

where $S(t)$ denotes the (even, entire) function

$$S(t) = \int_0^t (\sin v)^2/v \, dv. \tag{10}$$

In fact, if $\lambda(u)$ is defined to be $\mu'(u)/u$, then it is clear that (3) and (7) together are equivalent to (8), and that the relation (4), being identical with (6) by virtue of (5), appears in the form

$$F_\phi(t) = \exp\{-ct^2 - \int_{+0}^\infty \lambda(u)(\sin tu)^2/u \, du\}. \tag{11}$$

But it is clear from (10) that $dS(tu)/du = (\sin tu)^2/u$. Hence, a partial integration shows that the integral occurring in (11) is identical with

$$0 - 0 - \int_{+0}^\infty S(tu)\, d\lambda(u),$$

since, in view of (8) and (10), the integrated part, $\lambda(u)S(tu)$, tends to 0 whether $u \to \infty$ or $u \to 0$ (while t is fixed). Since this means that (11) is identical with (9), the assertion of (ii) follows.

(iii) *Every symmetric L-distribution ϕ is absolutely continuous for $0 < x < \infty$ (hence, by (2), for $-\infty < x < 0$). In order that ϕ be absolutely continuous for $-\infty < x < \infty$ (that is, in order that the jump $\phi(+0) - \phi(-0)$ be 0), it is necessary and sufficient that the data c, $\lambda(u)$ determining $\phi = \phi_c^\lambda$ satisfy the following alternative: Either $c > 0$ (while $\lambda(u)$ is arbitrary) or $c = 0$ but*

$$\int_{+0}^\infty \lambda(u)/u \, du = \infty. \tag{12}$$

This can be concluded from (i) and (ii), as follows: In view of (1) and of the Riemann-Lebesgue lemma, the first assertion of (iii) implies that $F_\phi(t) \to \phi(+0) - \phi(-0)$ as $t \to \infty$. Hence (8) and (9), where $c \geqq 0$, show that $\phi(+0) = \phi(-0)$ if and only if either $c > 0$ (while $\lambda(u)$ is arbitrary) or $c = 0$ but

$$\int_{+0}^\infty \lambda(u)(\sin tu)^2/u \, du \to \infty \text{ as } t \to \infty. \tag{13}$$

Since it is readily seen from (8) that (13) is equivalent to (12), this proves (iii).

3. With reference to any positive constant q, let $\phi_q = \phi_q(x)$ denote the symmetric L-distribution ϕ_c^λ for which c is 0 and $\lambda(u)$ is q or 0 according as $u \leqq 1$ or $u > 1$. Then (8) and (12) are satisfied, and (9) reduces to

$$(14) \qquad F_{\phi_q}(t) = e^{-qS(t)}, \text{ where } S(t) = \int_0^t (\sin u)^2/u\, du,$$

by (10). Hence (ii) and (iii) imply the first of the following three assertions:

(iv) *If q is a positive constant, then the Fourier inversion of the case (14) of (1) defines a distribution $\phi(x) = \phi_q(x)$ which is absolutely continuous for $-\infty < x < \infty$* (and is symmetric in the sense of (2), since (14) is an even function). *The density $\phi_q'(x) = \phi_q'(-x)$, where $\phi_q' = d\phi_q/dx$, is a monotone function of $|x|$* (which implies that $\phi_q'(\pm\infty) = 0$, and that $\phi_q'(+0) = \phi_q'(-0)$ exists if it is allowed to be ∞; actually, it will not or will be ∞ according as q does not or does exceed a certain critical value q_0, which will be determined in (vi) below). *For varying q, the distributions ϕ_q, where $0 < q < \infty$, form under convolution a semi-group which is cyclic in q:*

$$(15) \qquad \phi_{q_1+q_2} = \phi_{q_1} * \phi_{q_2},$$

and $\phi_q(x)$ tends, as $q \to +0$, to the (discontinuous) distribution

$$(16) \qquad \phi_0(x) = \tfrac{1}{2} + \tfrac{1}{2}\operatorname{sgn} x.$$

The second assertion of (iv) follows from (i) (it is understood that, when the density $\phi_q'(x)$ is claimed to *be* monotone for $0 < x < \infty$, what is actually meant is that $\phi_q'(x)$ *can be chosen* to be monotone for $0 < x < \infty$, since the density of an absolutely continuous distribution function is defined almost everywhere only; a corresponding understanding will hold when $\phi_q'(x)$ will be claimed to be continuous in (vi) below). Finally, since (15) is equivalent to the statement that the transform $F_q(t)$ belonging to $q = q_1 + q_2$ is the product of the transforms belonging to $q = q_1$ and $q = q_2$, and since the transform (1) of (16) is $F_{\phi_0}(t) \equiv 1$, the truth of (15) and of $\phi_{+0} = \phi_0$ is clear from (14).

(v) *A distribution is a symmetric L-distribution if and only if it is a limit, as $n \to \infty$, of distributions of the form*

$$(17) \qquad \phi(x) = \gamma_h(x) * \phi_{q_1}(p_1 x) * \cdots * \phi_{q_n}(p_n x),$$

where $\gamma_h(x)$ is the symmetric normal distribution of standard deviation

h ($\geqq 0$), the distribution $\phi_q(x)$ is defined by (14), *and $q_k = q_k(n) \geqq 0$ and $p_k = p_k(n) > 0$, where $k = 1, \cdots, n$, are $2n$ unspecified constants.*

This follows from (ii) and from Lévy's compactness theorem of the transform (1) of arbitrary distributions ϕ. In fact, it is readily seen from (10) that a given function $F_\phi(t)$ of t (where $-\infty < t < \infty$) is representable in the form (9), with some $c \geqq 0$ and with some $\lambda(u)$ satisfying (8), if and only if $F_\phi(t)$ is a limit (a limit which is uniform on *every fixed* finite t-interval) of functions of the form

$$F_\phi(t) = \exp(-ct^2) \prod_{k=1}^{n} \exp\{-q_k S(t/p_k)\}. \tag{18}$$

But it is clear from (14) and from the definition of γ_h (where $h \geqq 0$, and where (16) is meant to be the case $h = 0$ of γ_h) that (18) is equivalent to (17), where $c = c(h) \geqq 0$ according as $h \geqq 0$.

According to (v), the symemtric L-distributions (and only these) can be reduced to the semi-group ($0 < q < \infty$) of the particular distributions ϕ_q defined in (iv), if these distributions, after the adjunction of the symmetric normal distributions, are extended by the following three operations: changing of the unit of length on the x-axis, the convolution operation ($*$), and operation of a limit process ($n \to \infty$). This is the reason why the distributions ϕ_q, defined in (iv), will now be investigated in detail.

4. If a distribution $\phi(x)$ is absolutely continuous for $-\infty < x < \infty$, then the Fourier inversion of (1) is

$$2\pi\phi'(x) = \int_{-\infty}^{\infty} e^{-ixt} F_\phi(t)\, dt, \qquad \left(\int_{-\infty}^{\infty} = \lim_{T\to\infty} \int_{-T}^{T}\right), \tag{19}$$

valid at all those points x (if any) at which the density $\phi' = d\phi/dx$ satisfies a certain local condition. Such a condition is the monotony of ϕ' in some neighborhood of x (if $\phi'(x)$ is meant to be $\frac{1}{2}\phi'(x+0) + \phi'(x-0)$). It follows therefore from the first and from the second of the assertions of (iv) that (with the preceding parenthetical proviso) the representation (19) of the density of $\phi(x)$ is valid for every $\phi_q(x)$ at every $x \neq 0$. In view of (14), this means that

$$\pi\phi_q'(x) = \int_0^\infty e^{-qS(t)} \cos xt\, dt, \text{ where } S(t) = \int_0^t (\sin u)^2/u\, du, \tag{20}$$

whenever

$$0 < q < \infty \tag{21}$$

and

(22) $$0 < x < \infty$$

(and (22) can be replaced by $-\infty < x < \infty$, since

(23) $$\phi_q'(-x) = \phi_q'(x),$$

by (2)). But the point $x = 0$ is excluded in this deduction.

A partial integration of (20) shows that, under the assumptions (21) and (22),

(24) $$\pi\phi_q'(x)x/q = \int_0^\infty e^{-qS(t)}(\sin t)^2/t \sin xt\, dt.$$

Since it is clear from (10) that $0 \leqq e^{-S(t)} \leqq t^{-C}$, where $C > 0$ is independent of $t > 0$, it is clear that, for every fixed $q > 0$, the integral (24) is uniformly convergent for $-\infty < x < \infty$. It follows therefore from (24), and from the fact that (23) is monotone (non-increasing) on (22), that the product $\phi_q'(x)x$ (when defined at $x = 0$ as its limit for $x \to 0$) is continuous for $-\infty < x < \infty$. Hence, if $q > 0$ is arbitrary, $\phi_q'(x)$ is continuous on (22).

5. It also follows that there exists a constant $c = c_q$ satisfying

(25) $$\phi_q'(x) \sim c_q/x \text{ as } x \to 0, \text{ where } c_q \geqq 0$$

(it is understood that (25) means

(26) $$x\phi_q'(x) \to 0 \text{ as } x \to 0$$

if $c_q = 0$). It will now be shown that $c_q > 0$ or $c_q = 0$ according as q is or is not less than a certain critical value (which turns out to be the value $q = 2$).

It is seen from (10) that, as $t \to \infty$,

$$S(t) \sim \text{const.} \int_1^t du/u, \text{ where const.} = \pi^{-1}\int_0^\pi \sin^2 u\, du = \tfrac{1}{2},$$

hence $S(t) \sim \frac{1}{2}\log t$, and that this can be refined to the existence of a constant C having the property that

(27) $$S(t) - \tfrac{1}{2}\log t \to C \text{ as } t \to \infty$$

(in fact, the existence of such a constant follows from $\Sigma 1/n^2 < \infty$ in the same way as the existence of the Euler-Mascheroni constant). It is clear from (27) that, for every fixed $q > 0$,

(28) $$e^{-qS(t)} \sim Q/t^{\frac{1}{2}q} \text{ as } t \to \infty, \text{ where } Q = Q(q) = e^{-qC} > 0.$$

If $q > 2$, then it follows from (28) that, as $t \to \infty$, the integrand of (24) is majorized, uniformly for $-\infty < x < \infty$, by a constant multiple of $1/t^{1+\epsilon}$, where $\epsilon = \epsilon(q) > 0$. It follows therefore from (24) and (23) that, if $q > 2$, the density $\phi_q'(x)$ remains continuous at $x = 0$ and the graph of $y = y(x)$, where $y(x) = \phi_q'(x) = y(-x)$ and $-\infty < x < \infty$, will have at $x = 0$ a tangent parallel to the x-axis. If, on the other hand, $0 < q \leqq 2$, then, since (20) and (28) show that $\phi_q'(x)$ behaves, as $x \to \pm 0$, in the same way as

$$(29) \qquad \int_{+0}^{\infty} t^{-\frac{1}{2}q} \cos xt \, dt,$$

the density will become infinite at $x = 0$, the order of infinity being that of $\log |x|^{-1}$ or of $|x|^{-p}$, where $p = p(q) > 0$, according as $q = 2$ or $0 < q < 2$. Accordingly, the situation is as follows:

(vi) *If q (> 0) is arbitrary, then the density $\phi_q'(x)$ is monotone and continuous for $0 < x < \infty$ (hence, by (23), for $-\infty < x < 0$). If $q > 2$, then $\phi_q'(x)$ is continuous at $x = 0$ also. If $0 < q \leqq 2$, then $\phi_q'(x) \to \infty$ as $x \to \pm 0$.*

In fact, the first assertion of (vi), that concerning any $q > 0$, was the result of Section 4.

6. The second assertion of (vi) can be generalized as follows:

(vii) *If $q > 2m$, where m is a fixed positive integer, then $\phi_q(x)$ has a continuous $(m-1)$-st derivative for $-\infty < x < \infty$.*

In fact, the $(m-1)$-st derivative of the integral is majorized, uniformly for $-\infty < x < \infty$, by a constant multiple of

$$(30) \qquad \int_{0}^{\infty} t^{m-1} e^{-qS(t)} \, dt,$$

and (28) shows that (30) is a convergent integral if $m - 1 - \frac{1}{2}q > -1$, that is, if $q > 2m$.

(viii) *If $0 < x < \infty$ (or $-\infty < x < 0$), then, whenever $q > 0$, there exists a (continuous) n-th derivative $d^n \phi_q(x)/dx^n$ for $n = 1, 2, \cdots$.*

For $n = 1$, this was proved (Section 4) by a partial integration, that leading from (20) to (24). For an arbitrary n, the assertion of (viii)

follows, after n partial integrations, in the same way as it did for $n=1$. Actually, (viii) is not the last word, since every $\phi_q(x)$ turns out to be regular (analytic) at every real $x \neq 0$.

There is a general principle concerning arbitrary distributions ϕ, which, roughly, is to the effect that "properties of smoothness," possessed by $\phi(x)$ for $-\infty < x < \infty$, are not decreased, and have a tendency to be increased, if $\phi(x)$ is replaced by $\phi(x) * \psi(x)$, where $\psi(x)$ is any distribution. If this is applied to (15), what results is that, if $0 < q < r$, then ϕ_q is at least as "smooth" as ϕ_r (and perhaps "smoother"). Clearly, (vi) and (vii) contain precise formulations of this expectation. On the other hand, (viii) is a manifestation of the general principle only insofar as the C^∞-character of $\phi_q(x)$ on (22) is claimed for *every* value q (large *or* small); so that the "smoothing effect" (when ϕ_q is replaced by ϕ_r, where $r > q$) takes hold only at $x=0$, at the point excluded in (viii). This remark will be essential in what follows.

7. Theorems (iv) and (vi)-(vii) are purely analytical in nature; they could hardly be expected from the rôle which, according to (v), the semi-group (15), defined on (4) by the Fourier inversion of the case (14) of (1), plays in the theory of symmetric L-distibutions. As a matter of fact, all assertions corresponding to (iv) and (vi)-(vii) become false for the semi-group of distributions, say ψ_q, which belong to the class of all symmetric, infinitely divisible distributions in the same way as the distributions ϕ_q, discussed above, belong to the more restricted class of all symmetric L-distributions. This will be shown by a discussion of the distributions ψ_q just indicated.

First, if the class of all symmetric L-distributions $\phi(x)$ is replaced by the larger class of all symmetric, infinitely divisible distributions, then (11) and (8) become replaced by (6) and (3). Hence it is clear that the distributions ϕ_q, defined for every $q > 0$ by (14), must be replaced by the distributions ψ_q which, for every $q > 0$, are defined by

$$(31) \qquad F_{\psi_q}(t) = e^{-qR(t)}, \text{ where } R(t) = \int_0^{|t|} (\sin u / u)^2 \, du.$$

In fact, it is clear that (4) reduces to (31) if the arbitrary $\mu(u)$, which is subject only to (3), is chosen as follows: $\mu(u) = 0$ or $\mu(u) = q$ according as $0 \leqq u \leqq q$ or $q < u < \infty$. Correspondingly, it is clear that (v) becomes true

for the extended ϕ-class if ϕ_q is replaced by ψ_q in (17) (incidentally, what corresponds to (12) in the appropriate analogue of (iii) is

$$(32) \qquad \int_{+0}^{\infty} u^{-2}\, d\mu(u) = \infty\,;$$

cf. [25], Theorem (III), p. 289, formula (8), p. 286, and Corollary, p. 288). Finally, it is clear from (31) that, corresponding to (15),

$$(33) \qquad \psi_{q_1} * \psi_{q_2} = \psi_{q_1+q_2}$$

and that $\psi_q \to \psi_0$ as $q \to 0$, if ψ_0 denotes the distribution which in (16) is denoted by ϕ_0.

There is however a fundamental difference between ϕ_q and ψ_q (for any fixed $q > 0$). In fact, such results as (i) or (iv) depend on (27) and so, in particular, on the circumstance that $S(\infty) = \infty$ in (14). In contrast, $R(\infty) < \infty$ in (31), since

$$(34) \qquad R(t) = \int_0^t (\sin u/u)^2\, du \to \tfrac{1}{2}\pi \text{ as } t \to \infty.$$

A first implication of this contrast is that the jump $\phi(+0) - \phi(-0)$, which was 0 for $\phi = \phi_q$, is $e^{-\frac{1}{2}\pi q} > 0$ for $\phi = \psi_q$; in fact, $F_{\psi_q}(t) \to e^{-\frac{1}{2}\pi q}$ as $t \to \infty$, by (31) and (34). This contrast alone would not be serious, since the jump can be disposed of by first subtracting from $\psi_q(x)$ the function $e^{-\frac{1}{2}\pi q}\phi_0(x)$, where $\phi_0(x)$ is defined by (16), and then dividing the resulting monotone function

$$(35) \qquad \theta_q(x) = \psi_q(x) - e^{-\frac{1}{2}\pi q}\phi_0(x)$$

by the (positive) constant $\psi_q(\infty) - e^{-\frac{1}{2}\pi q}\phi_0(\infty) = 1 - e^{-\frac{1}{2}\pi q}$. In order to simplify the formulae, this trivial re-normalization will be disregarded, that is, the function (35) (which differs from a distribution in a positive constant factor) will be used as it stands. Since $F_{\phi_0}(t) \equiv 1$, by (1) and (16), it follows from (31), (35) and (1) that

$$(36) \qquad F_{\theta_q}(t) = F_{\psi_q}(t) - F_{\psi_q}(\infty), \text{ where } F_{\psi_q}(\infty) = e^{-\frac{1}{2}\pi q}.$$

Clearly, (33) becomes a relation connecting the three functions $\theta_q(x)$ belonging to $q = q_1, q_2, q_1 + q_2$, if ψ_q is substituted from (35) in terms of θ_q. But it turns out that (31), in contrast to (15), fails to induce an "increase in smoothness" (cf. the end of Section 6) when q increases.

8. This can be seen as follows:

According to (34), there exists a positive constant C for which

$$\tfrac{1}{2}\pi - R(t) \sim C/t \text{ as } t \to \infty \qquad (C > 0) \tag{37}$$

(this C is not the same as, and will play a more delicate part than, the C occurring in (27)). What corresponds to (28) is that, for every fixed $q > 0$,

$$F_{\theta_q}(t) \sim q/t \text{ as } t \to \infty, \text{ where } q = Q(q) = Cqe^{-\frac{1}{2}\pi q} > 0, \tag{38}$$

as seen from (31), (36) and (37).

Since (38) and the relation $F_{\theta_q}(-t) = F_{\theta_q}(t)$ imply that the integral (over the entire t-region $-\infty < t < \infty$) of the square of $F_{\theta_q}(t)$ is finite, it follows from Plancherel's theorem that the case $\phi = \theta_q$ of (1) belongs to an absolutely continuous $\phi(x)$ possessing a density $\phi'(x)$ the square of which has a finite integral over $-\infty < x < \infty$. In particular, $\theta_q(x)$ is absolutely continuous for $-\infty < x < \infty$, with a density $\theta_q'(x)$ which, corresponding to (19), is given by

$$\pi\theta_q'(x) = \int_0^\infty F_{\theta_q}(t) \cos xt \, dt, \tag{39}$$

as soon as any of the standard conditions for the validity of Fourier's theorem is assured. But the proof of (viii), which depends only on successive partial integrations, can be readily repeated when (14) is replaced by (31). Hence $\theta_q(x)$ has derivatives of arbitrarily high order for $0 < x < \infty$, and therefore for $-\infty < x < 0$. In particular, $\theta_q'(x)$ is differentiable at every $x \neq 0$, and so (39) is valid whenever (21) and (22) hold.

Finally, it is seen from (39) and (38) that, if q (> 0) is fixed, $\theta_q'(x) \to \infty$ as $x \to 0$, whether q be small or large. It follows therefore from (35) and (16) that the same is true if $\theta_q'(x)$ is replaced by $\psi_q'(x)$. Hence nothing like (vii) can hold if $\phi_q(x)$ is replaced by $\psi_q(x)$.

9. In order to complete the proof of the statements made in Section 7, it remains to be shown that the first part of the first assertion of (vi) breaks down if $\phi_q'(x) = \phi_q'(-x)$ is replaced by $\psi_q'(x) = \psi_q'(-x)$; in other words, that it is *not* true that all the distributions $\psi_q(x) = 1 - \psi_q(-x)$ are unimodal.

According to the analogue of (v), formulated in Section 6, every symmetric, infinitely divisible distribution $\phi(x)$ can be derived from the

particular distributions $\psi_q(x)$, belonging to the range (21), by certain trivial processes. According to [12], pp. 77-79, each of these processes preserves symmetric unimodal character.* Since every $\psi_q(x)$ is symmetric, it follows that, if every $\psi_q(x)$ were unimodal, then every symmetric, infinitely divisible $\phi(x)$ had to be unimodal and so, in particular, such as to possess no jump at any x distinct from $x=0$. But this is disproved by the example of the distribution $\phi(x)$ which is the convolution of $\pi_h(x)$ and $1-\pi_h(x)$ (cf. the end of Section 1), where $\pi_h(x)$ denotes Poisson's distribution of standard deviations h (>0).

APPENDIX IV.

Bell-shaped frequency curves mod 1.

Let the case $a=2$ of (1) be written in the form

$$\theta(q;\phi)=1+2\sum_{m=1}^{\infty}q^{m^2}\cos m\phi. \tag{1}$$

There are two (substantially different) fundamental properties of this θ: what results from the "linear transformations" property of the elliptic modular functions (a result the explicit form of which is

$$\theta(q;\phi)=\sum_{k=-\infty}^{\infty}(\pi/Q^2)^{-\frac{1}{2}}\exp\{-(2\pi k+\phi)^2/(2Q)^2\},\ 2\pi^2Q^2=-\log q, \tag{2}$$

the case $a=2$ of (5)-(6) in Appendix I), and the existence of an Eulerian factorization (in the sense of Hecke),

$$\theta(q;\phi)=\sum_{n=1}^{\infty}(1-q^{2n})(1+2q^{2n-1}\cos x+q^{4n-2}) \tag{3}$$

(Jacobi; cf. Pólya-Szegö, chap. I, no. 53). In (2), the arbitrary parameter

* The symmetry restriction is here essential, since the "rearrangement" lemma of Hardy and Littlewood, used in [12], pp. 77-79, cannot be generalized to the statement that the convolution of two arbitrary (not necessarily symmetric) unimodal distributions is unimodal. This observation, made by Paul Lévy (cf. *Revue Mathématique de l'Union Interbalkanique*, vol. 2 (1939), p. 22), was recently rediscovered by K. L. Chung (*Comptes Rendus*, vol. 236 (1953), pp. 583-584, and [22], pp. 254-255). Chung's counterexample, which he needs in order to disprove the validity of a proof given by Gnedenko and Kolmogoroff (Lapin; cf. [22], p. 160), is substantially the same as Lévy's. The non-existence of an extension of [12], pp. 78-79, seemed to be obvious when [12] was written.

Q (>0) corresponds to the choice of the standard deviation of a symmetric normal distribution, a *non-cyclic* distribution $(-\infty<\phi<\infty)$ the density of which is the 0-th term $(k=0)$ of the series (2).

If $f(\phi)$ is any real-valued (and, for the sake of simplicity, continuous) function of period 2π, let $f(\phi)$ be called *unimodal* $(\bmod 2\pi)$ if, in the cyclic interpretation of the (ϕ,f)-diagram, the graph of $f=f(\phi)$ consists of a single convex and of a single concave arc $(\bmod 2\pi)$. If $f(\phi)$ has a continuous second derivative $f''(\phi)$, this means that either $f(\phi)$ is a constant or $f''(\phi)$ changes sign (at most and/or exactly) *twice* within a period. If $f(\phi)=f(-\phi)$, this will be referred to as the *bell-shaped* character $(\bmod 2\pi)$ of $f(\phi)$.

In this terminology, the second (that is, the stronger) of the two properties, referred to in the last but one paragraph of Appendix I, can be formulated as follows: *$f(\phi)=\theta(q;\phi)$ is bell-shaped for every fixed positive $q<1$.* This was proved in [18] (it seems to be surprising that a proof was not given many years earlier; an explanation may be that the issue was made acute only by the rôle of (1) as the density of the "cyclic normal distribution," that is, by the interepretation of (2) as the reduction $\bmod 2\pi$ of the linear "normal frequency law"). If Hilfssatz I of Pólya [31], p. 126, is applied to the second derivative of (1) with respect to ϕ, then the difficulty in proving the bell-shaped character $(\bmod 2\pi)$ of (1) for a fixed q (>0) is seen to be a *monotone* increasing function of q (<1).

The proof in [18] is not a pretty one: it starts out with (2) but, since (2) in itself cannot succeed,* the main point in the proof depends on an appeal to Rouché's theorem, which seems to be entirely out of place in this "explicit" context. Subsequently, the result of [18] was therefore verified by Koschmieder along "explicit" lines; cf. [29] and [30], where the literature of earlier attempts will be found. But Koschmieder's proof represents the other extreme, since it involves an impressive cross-section of the elliptic *Formelapparat*.

Since (3) leads, cf. [17], to quite a trivial proof of the weaker statement

* As a matter of fact, (2) and the mere circumstance that $g(x)=\exp(-cx^2)$ is a monotone function of $|x|$ if $c>0$ (or, for that matter, that this $g(x)$ is bell-shaped for $-\infty<x<\infty$) fail to imply even what in the last but one paragraph of Appendix I was referred to as the *weaker* property. For it clearly is not true that if $g(x)>0$ is an unspecified even function which is monotone for $0\leqq x<\infty$ (or, for that matter, which is bell-shaped for $-\infty<x<\infty$) and tends to 0 with a sufficient rapidity as $x\to\infty$, then the function $f(\phi)=\sum g(\phi+2\pi k)=f(\phi+2\pi)$, where $k=\cdots,-1,0,1,\cdots$, must be monotone on the interval $0<\phi<\pi$.

that (1) is monotone for $0 \leqq x \leqq \pi$ (at every fixed positive $q < 1$), it was natural to try a similar approach to the stronger statement, that claiming the bell-shaped character (mod 2π) of (1). The result turns out to be unexpected, since it shows that the matter has nothing to do with elliptic modular functions as such. In fact, the following lemma will now be proved:

If $c_1, c_2, \cdots$ is any sequence of numbers satisfying

(4) $$0 \leqq c_n \leqq 1 \text{ and } \Sigma c_n < \infty,$$

then the function

(5) $$f(\phi) = \Pi f_n(\phi),$$

where

(6) $$f_n(\phi) = 1 + 2c_n \cos \phi + c_n^2,$$

is bell-shaped (mod 2π).

It is clear from (3) that the result of [18] on (1) follows from this lemma by choosing $c_n = q^{2n-1}$; in fact, the factor $\Pi(1 - q^{2n}) = \text{const.} > 0$ is immaterial. Similarly, (6) shows that (5) can be written in the form

$$f(\phi) = C\,\Pi(1 + a_n \cos \phi), \qquad a_n = 2c_n/(1 + c_n^2),$$

where the factor $C = \Pi(1 + c_n^2)$ is immaterial. Since the conditions (4) remain unaltered if every c_n is replaced by the corresponding a_n, it is seen that the lemma can be formulated as follows:

If $a_1, a_2, \cdots$ is a sequence of numbers satisfying $0 \leqq a_n \leqq 1$ and $\Sigma a_n < \infty$, then the function $\Pi(1 + a_n \cos \phi)$ is bell-shaped (mod 2π).

For reasons of continuity, it is sufficient to prove the lemma for the case in which the first two of the three conditions (4) are replaced by $0 < c_n < 1$. Then the functions (6) and (5) are positive throughout. Two differentiations of (6) and (5) show that

(7) $$-\tfrac{1}{2}f''(\phi)/f(\phi) = 2(F_2 - F_1^2) \sin^2 \phi + F_1 \cos \phi,$$

if $' = d/d\phi$, $F_1^2 = (F_1)^2$, and

(8) $$F_j = F_j(\phi) = \Sigma(c_n/f_n)^j, \qquad f_n = f_n(\phi) = f_n(-\phi),$$

where $j = 1$ or $j = 2$ (eventually, (8) will be needed for $j = 3$ also.) The simple structure of (7) and (8) will lead to a straightforward proof.

It can be assumed that ϕ is confined to the interval $0 < \phi < \pi$. Then

the assertion, to be proved, is that $f''(\phi)$ does not change sign more than once. It is clear that if $t(\phi)$ denotes the trigometric function $\frac{1}{2}\cos\phi/\sin^2\phi$, then (7) shows that $f''(\phi)$ will vanish when and only when the function

$$F_1 - F_2/F_1 \tag{9}$$

of ϕ becomes equal to $t(\phi)$. But the derivative of $t(\phi)$ is seen to be negative (since $0 < \phi < \pi$), and so $t(\phi)$ is decreasing throughout. Hence the proof will be complete if it is shown that the function (9) of ϕ is increasing throughout.

Differentiation of (5) and (8) shows that $F_1' = 2F_2 \sin\phi$ and $F_2' = 4F_3 \sin\phi$. Hence the derivative of (9) is seen to be of the same sign as

$$(F_1^2F_2 + F_2^2 - 2F_1F_3)\sin\phi.$$

Since $\sin\phi > 0$ (for $0 < \phi < \pi$), it follows that the function (9) is increasing throughout if it is true that

$$2F_1F_3 < F_1^2F_2 + F_2^2. \tag{10}$$

Let ϕ be fixed. Then every sum (8) is of the form

$$F_j = \sum_{n=1}^{\infty} p_n{}^j, \text{ where } p_n > 0, \tag{11}$$

since $c_n > 0$ and $f_n(\phi) > 0$. But (11) implies that $F_2^2 \leqq F_1F_3$ (Schwarz). Hence (10) is certainly true if $2F_3 < F_1F_2 + F_3$, that is, if

$$F_3 < F_1F_2. \tag{12}$$

Finally, since $\Sigma\Sigma\, p_np_m{}^2 > \Sigma\, p_np_n{}^2$ if every p is positive, (12) is clear from (11).

THE JOHNS HOPKINS UNIVERSITY.

REFERENCES.

[1] H. Bateman, *Higher transcendental functions*, vol. 3, chap. XVIII (pp. 206-227), New York, 1955.

[2] G. H. Hardy, *Ramanujan* (chap. XI, pp. 186-211), Cambridge, 1940.

[3] ——— and J. E. Littlewood, "Contributions to the theory of the Riemann zeta-function and the theory of the distribution of primes," *Acta Mathematica*, vol. 41 (1918), pp. 119-196.

[4] P. Hartman and A. Wintner, "On the spherical approach to the normal distribution law," *American Journal of Mathematics*, vol. 62 (1940), pp. 759-779.

[5] I. I. Hirschman and D. V. Widder, *The convolution transform*, Princeton, 1955.

[6] P. Lévy, *Théorie de l'addition des variables aléatoires*, Paris, 1937.

[7] E. Lindelöf, *Le calcul des résidues* (chap. V, pp. 108-141), Paris, 1905.

[8] H. Mellin, "Die Dirichletschen Reihen, die zahlentheoretischen Funktionen und die unendlichen Produkte von endlichem Geschlecht," *Acta Societatis Scientiarum Fennicae*, vol. 31 (1903), no. 2.

[9] R. E. A. C. Paley and N. Wiener, *Fourier transforms in the complex domain*, New York, 1934.

[10] H. Pollard, "The representation of $\exp(-x^{\lambda})$ as a Laplace integral," *Bulletin of the American Mathematical Society*, vol. 52 (1946), pp. 908-910, and "The completely monotonic character of the Mittag-Leffler function $E_a(-x)$," *ibid.*, vol. 54 (1948), pp. 1115-1116.

[11] C. R. Putnam and A. Wintner, "On the addition of symmetric normal frequency curves," *Mathematicae Notae* (Rosario), vol. 11 (1952), pp. 79-86.

[12] A. Wintner, "On a class of Fourier transforms," *American Journal of Mathematics*, vol. 58 (1936), pp. 45-90.

[13] ———, "The singularities of Cauchy's distributions," *Duke Mathematical Journal*, vol. 8 (1941), pp. 678-681.

[14] ———, "The sum formula of Euler-Maclaurin and the inversions of Fourier and Möbius," *American Journal of Mathematics*, vol. 69 (1947), pp. 685-708.

Appendix I.

[15] P. Lévy, "L'addition des variables aléatoires définies sur une circonference," *Bulletin de la Société Mathématique de France*, vol. 67 (1939), pp. 1-41. Cf. also [23], pp. 186-191.

[16] A. Wintner, "On the stable distribution laws," *American Journal of Mathematics*, vol. 55 (1933), pp. 335-339.

[17] ———, "On the shape of the angular case of Cauchy's distribution curves," *Annals of Mathematical Statistics*, vol. 18 (1947), pp. 589-593.

[18] ———, "The real elliptic ϑ_3-function," *Duke Mathematical Journal*, vol. 15 (1948), pp. 407-411.

[19] ———, "On a decomposition into singularities of theta-functions of fractional index," *American Journal of Mathematics*, vol. 71 (1949), pp. 105-108.

[20] F. Zernike, "Wahrscheinlichkeitsrechnung und mathematische Statistik," *Handbuch der Physik*, vol. 3 (1928), pp. 419-492.

Appendix II.

[21] M. Girault, "Les fonctions charactéristiques et leur transformations," *Publications de l'Institut de Statistique de l'Université de Paris*, vol. 4 (1955), pp. 223-298.

[22] B. V. Gnedenko and A. N. Kolmogorov, *Limit distributions for sums of independent random variables*, Cambridge, Mass., 1954.

[23] P. Lévy, *Processus stochastiques et mouvement Brownien*, Paris, 1948.

[24] A. Wintner, "Sur les distribution symétrique à fonctions caractéristiques convexes," *Publications de l'Institut de Statistiques de l'Université de Paris* (to appear).

Appendix III.

[25] P. Hartman and A. Wintner, "On the infinitesimal generators of integral convolutions," *American Journal of Mathematics,* vol. 64 (1942), pp. 273-298.

[26] E. Hille, *Functional analysis and semi-groups,* New York, 1948.

[27] P. Lévy, "Sur les intégrales dont les éléments sont des variables aléatoires indépendentes," *Annali della R. Scuola Normale Superiore di Pisa,* ser. 2, vol. 3 (1934), pp. 337-366; cf. *ibid.,* vol. 4 (1935), pp. 217-218.

[28] J. von Neumann and I. J. Schoenberg, "Fourier integrals and metric geometry," *Transactions of the American Mathematical Society,* vol. 50 (1941), pp. 226-251.

Appendix IV.

[29] L. Koschmieder, "Die Krümmung des Schaubildes der Jacobischen Etafunktion," *Anzeiger der Oesterreichischen Akademie der Wissenschaften,* 1948, pp. 210-214.

[30] ———, "Die Krümmung des Schaubildes der Jacobischen Thetafunktion," *Revista Hispano-Americana,* ser. 4, vol. 8 (1948), pp. 257-260.

[31] G. Pólya, "Qualitatives über Wärmeausgleich," *Zeitschrift für angewandte Mathematik und Mechanik,* vol. 13 (1933), pp. 125-128.

THE DECOMPOSITION THEOREM FOR V-MANIFOLDS.* [1]

By Walter L. Baily, Jr.

1. Introduction. The purpose of this paper is to give a more or less self-contained development of the analytical tools employed in the theory of harmonic integrals in such a way that the theorems we obtain may be applied to fairly general situations. More specifically we wish to present in fairly general context a proof of the so-called regularity theorem as well as a proof of the orthogonal decomposition theorem for compact V-manifolds in a form that will apply for C^∞ forms with coefficients in a vector bundle. In this way we desire to complement a paper of Satake [9] by obtaining Hodge's theorem for compact V-manifolds, present an analytical background for recent papers of Nakano [7], and lay the essential analytical foundations for a forthcoming paper of the author. In our proof we lay no claim to originality but simply make certain modifications in a proof presented by Kodaira in a course given in 1952-53 which in turn represented a certain modification of an earlier proof, also by Kodaira [4], which had been based on methods of Hadamard.

The author wishes to express his thanks to K. Kodaira for many helpful suggestions and conversations. He also wishes to thank I. Satake for the communication of the unpublished manuscript of his paper in which the notion of V-manifold is introduced, and in which an analog of de Rham's theorem is proved for V-manifolds.

2. V-manifolds and vector bundles. We recall here briefly the definition of V-manifold due to Satake from whose paper we carry over certain definitions for reference. By a local uniformizing (l. u. s.) $\{U, G, \varphi\}$ for an open subset $\tilde{U}$ of a Hausdroff space $\mathfrak{V}$ we mean a collection of the following objects:

U: a connected open neighborhood of the origin in R^n.

G: a finite group of C^∞ transformations of U.

* Received September 23, 1955; revised April 10, 1956.

[1] The author wishes to acknowledge that the final correction of manuscript for this paper was carried out while the author was employed on a project supported by a National Science Foundation Grant.

φ: a continuous map of U onto $\tilde{U}$, such that $\varphi \circ \sigma = \varphi$ for all $\sigma \in G$ and such that the induced map of U/G onto U is a homeomorphism. (Here $\circ$ denotes composition of functions.)

Let $\{U, G, \varphi\}$ and $\{U', G', \varphi'\}$ be l. u. s.'s for $\tilde{U}$ and $\tilde{U}'$ respectively such that $\tilde{U} \subset \tilde{U}'$; by an injection of $\{U, G, \varphi\}$ into $\{U', G', \varphi'\}$ we mean a C^∞ isomorphism λ of U into U' such that for any $\sigma \in G$ there exists $\sigma' \in G'$ satisfying the relations $\lambda \circ \sigma = \sigma' \circ \lambda$ and $\varphi = \varphi' \circ \lambda$. Then a C^∞ V-manifold shall consist of a connected Hausdorff space $\mathfrak{V}$ and a family $\mathfrak{F}$ of l. u. s.'s for open subsets of $\mathfrak{V}$ satisfying the following conditions:

(i) If $\{U, G, \varphi\}$, $\{U', G', \varphi'\} \in \mathfrak{F}$ and $\tilde{U} = \varphi(U)$ is contained in $\tilde{U}' = \varphi'(U')$, then there exists an injection of $\{U, G, \varphi\}$ into $\{U', G', \varphi'\}$.

(ii) The open sets $\tilde{U}$, for which there exists a l. u. s. $\{U, G, \varphi\} \in \mathfrak{F}$, form a basis of open sets in $\mathfrak{V}$.

Real and complex V-manifolds are defined in a similar fashion. Moreover, a V-manifold is orientable if all the injections in (i) are orientation preserving and if for any l. u. s. $\{U, G, \varphi\}$ the action on U of each $\sigma \in G$ is orientation-preserving.

If $\mathfrak{O}$ is an open subset of $\mathfrak{V}$, we denote by $\mathfrak{F}_{\mathfrak{O}}$ the subfamily of $\mathfrak{F}$ consisting of those l. u. s. $\{U, G, \varphi\}$ such that $\tilde{U} \subset \mathfrak{O}$. By a differential form ϕ on $\mathfrak{O}$ we mean a collection of differential form $\{\phi_U\}$, where ϕ_U is a differential form on U invariant under G for $\{U, G, \varphi\} \in \mathfrak{F}_{\mathfrak{O}}$, such that if $\lambda: \{U', G', \varphi'\} \to \{U, G, \varphi\}$ is an injection, $\lambda^* \phi_U = \phi_{U'}$ where λ^* is the mapping of differential forms dual to λ.

By a V-bundle B over $\mathfrak{V}$ with group Γ and fibre F we mean that there is given for each l. u. s. $\{U, G, \varphi\} \in \mathfrak{F}$ a bundle B_U over U with group Γ and fibre F together with an anti-isomorphism h_U of G into a group of bundle maps of B_U onto itself such that if b lies in the fibre over $x \in U$, then $h_U(g)b$ lies in the fibre over $g^{-1}x$ for $g \in G$; and moreover, if λ is an injection, $\lambda: \{U, G, \varphi\} \to \{U', G', \varphi'\}$, then we are given a bundle map $\lambda^*: B_{U'} \mid \lambda(U) \to B_U$ satisfying the requirements that if $g \in G$ and g' is the unique element of G' such that $\lambda \circ g = g' \circ \lambda$, then $h_U(g) \circ \lambda^* = \lambda^* \circ h_{U'}(g')$, and that if $\lambda\{U, G, \varphi\} \to \{U', G', \varphi'\}$ and $\lambda': \{U', G', \varphi'\} \to \{U'', G'', \varphi''\}$ are injections, then $(\lambda \circ \lambda')^* = \lambda'^* \circ \lambda^*$.

If $\mathfrak{V}$, Γ, F, and all B_U have a certain type of differentiable structure and if each λ^* and each $h_U(g)$, $g \in G$, is compatible with this type of differentiable structure, we say that the given V-bundle has this type of differentiable structure. In this paper we shall be concerned mainly with V-bundles

for which F is a real or complex vector space and Γ a real or complex linear group. Without loss of generality we may always suppose for a given V-bundle B over $\mathfrak{V}$ and for all l. u. s. $\{U, G, \varphi\} \in \mathfrak{F}$ that B_U is the product bundle over U; then if B is a V-bundle with q-dimensional vector space as fibre, $h_U(g)$ is a continuous mapping of U into the manifold of non-singular $q \times q$ matrices, for $g \in G$.

If B is a V-bundle over $\mathfrak{V}$ and $\mathfrak{O}$ an open subset of $\mathfrak{V}$, then by a section ϕ of B over $\mathfrak{O}$ we shall mean a collection of functions $\{\phi_U\}$, where ϕ_U is a function from U into B_U for each $\{U, G, \varphi\} \in \mathfrak{F}_{\mathfrak{O}}$ satisfying the following conditions:

s(i) If $x \in U$, $\phi_U(x)$ is in the fibre over x and $h_U(g)\phi_U(x) = \phi_U(g^{-1}x)$ for each $g \in G$.

s(ii) If $\lambda : \{U, G, \varphi\} \to \{U', G', \varphi'\}$ is an injection, then

$$\lambda^*\phi_{U'}(\lambda x) = \phi_U(x).$$

We speak of ϕ as being measurable, continuous, C^∞, etc., when all the functions ϕ_U are respectively measurable, continuous, C^∞, etc. Moreover, from this definition it is also obvious what we will mean by a C^∞ or analytic real or complex valued function on $\mathfrak{O}$. If, moreover, B is a vector bundle, then by a metric $\mathfrak{a}$ for B we shall mean a collection of functions $\{\mathfrak{a}_U\}$, where for each $\{U, G, \varphi\} \in \mathfrak{F}$ $\mathfrak{a}_U$ is a function which assigns to each $x \in U$ a (symmetric or hermitian) positive definite bilinear form $\mathfrak{a}_U(x)$ in the fibre of B_U over x such that the following conditions are fulfilled:

m(i) If α and β belong to the fibre of B_U over x, then

$$\mathfrak{a}_U(x)(\alpha, \beta) = \mathfrak{a}_U(g^{-1}x)(h_U(g)\alpha, h_U(g)\beta)$$

for $g \in G$.

m(ii) If $\lambda : \{U, G, \varphi\} \to \{U', G', \varphi'\}$ is an injection, then

$$\mathfrak{a}_U(x)(\lambda^*\alpha, \lambda^*\beta) = \mathfrak{a}_{U'}(\lambda x)(\alpha, \beta),$$

α and β being in the fibre of $B_{U'}$ over λx.

m(iii) If ψ and ϕ are continuous sections of B, the function $\mathfrak{a}(\psi, \phi)$ on $\mathfrak{V}$ defined by

$$\mathfrak{a}(\psi, \phi)(\varphi(x)) = \mathfrak{a}_U(x)(\psi_U(x), \phi_U(x)),\ x \in U,$$

is continuous.

If $\mathfrak{a}(\psi, \phi)$ is a C^∞ function on $\mathfrak{V}$ whenever ϕ and ψ are of class C^∞, then we say that $\mathfrak{a}$ is of class C^∞.

From now on we assume that $\mathcal{V}$ is paracompact. Let $\{\tilde{U}_i\}$ be a locally finite covering of $\mathcal{V}$ by open sets $\tilde{U}_i$ such that $\{U_i, G_i, \varphi_i\} \in \mathcal{F}$. By a C^∞ partition of unity subordinate to $\{\tilde{U}_i\}$, we mean a collection of C^∞ functions $\{l_i\}$ on $\mathcal{V}$ such that the support $\mathfrak{L}(l_i)$ of l_i is contained in $\tilde{U}_i$ and such that for each $y \in \mathcal{V}$, $\sum_i l_i(y) = 1$. It is easily seen that given a locally finite covering $\{\tilde{U}_i\}$ of $\mathcal{V}$, there exists a C^∞ partition of unity subordinate to $\{\tilde{U}_i\}$. We simply let $\{\tilde{V}_i\}$ be a shrinkage of $\{\tilde{U}_i\}$ and by an obvious averaging process on U_i with respect to G_i obtain for each i a C^∞ function m_i on $\mathcal{V}$ such that $m_i = 1$ on $\tilde{V}_i$ and $m_i = 0$ outside $\tilde{U}_i$; we then put $l_i = m_i/(\sum m_i)$.

We can now show that if B is a V-bundle over $\mathcal{V}$ with a vector space as fibre, then there exists a C^∞ metric for B. Let $\{U_i\}$ be a locally finite covering of $\mathcal{V}$ such that $\{U_i, G_i, \varphi_i\} \in \mathcal{F}$ and such that B_{U_i} is the product bundle, and let $\{l_i\}$ be a C^∞ partition of unity subordinate to $\{U_i\}$. Then $B_{U_i} = U_i \times V$, where V is a q-dimensional vector space. Let $v_1, \cdots, v_q$ be a fixed basis for V and let

$$b_{U_i}(x)\left(\sum c_k v_k, \sum d_k v_k\right) = \sum c_k \bar{d}_k.$$

If α and β belong to $y \times V$, $y \in U_i$, define

$$b_{U_i}{}^*(y)(\alpha, \beta) = l_i(\varphi_i(y))(\operatorname{ord} G_i)^{-1} \sum_{g \in G_i} b_{U_i}(g^{-1}y)(h_{U_i}(g)\alpha, h_{U_i}(g)\beta).$$

Finally define

$$a_{U_i}(x)(\alpha, \beta) = \sum_j b_{U_j}{}^*(x_{ij})(\lambda_{ij}\alpha, \lambda_{ij}\beta), \ x \in U_i,$$

where the summation extends over all j such that $\varphi_j^{-1} \circ \varphi_i(x)$ is not empty, where x_{ij} is a point of $\varphi_j^{-1} \circ \varphi_i(x)$, and where λ_{ij} is defined as follows: Let $\varphi_i(x) \in \tilde{U} \subset \tilde{U}_i \cap \tilde{U}_j$, $\{U, G, \varphi\} \in \mathcal{F}$, and let $\lambda_i^* : B_{U_i} \mid \lambda_i(U) \to B_U$ and $\lambda_j^* : B_{U_j} \mid \lambda_j(U) \to B_U$ be the bundle maps associated with injections λ_i and λ_j; then $\lambda_{ij} = \lambda_j^{*-1} \circ \lambda_i^*$. From the construction of $b_{U_i}{}^*$ it is clear that a_{U_i} is well-defined, that the collection of functions $\{a_{U_i}\}$ satisfy m(i) and m(ii), and that $a(\psi, \phi)$ is a function of class C^∞ if ψ and ϕ are of class C^∞. Finally a_U is uniquely defined by the requirement m(ii) for all $\tilde{U}$ such that $\{U, G, \varphi\} \in \mathcal{F}$, for the totality of all such $\tilde{U}$ is a base of neighborhoods for $\mathcal{V}$.

If B is a V-bundle over $\mathcal{V}$ with a finite dimensional vector space X as fibre, we can define the dual V-bundle B' of B. The fibre X' of B' is the dual space of X and the maps $h_U(g)'$ and $\lambda^{*\prime}$ for B' are defined to be the inverses of the transposes of $h_U(g)$ and λ^* respectively, the latter being the bundle maps previously defined for B. A C^∞ metric a for B gives rise to a

C^∞ metric $\mathfrak{a}'$ for B' in the following way: If $\{U, G, \varphi\} \in \mathfrak{F}$, then for each $x \in U$, $\mathfrak{a}_U(x)$ is a positive definite bilinear form in the fibre X_x of B_U over x. $\mathfrak{a}_U(x)$ gives rise naturally to a positive definite, bilinear form $\mathfrak{a}'_U(x)$ on the dual space X'_x, the fibre of B'_U over x, such that the matrical form of $\mathfrak{a}'_U(x)$ with respect to a fixed basis in X'_x is the transposed inverse of $\mathfrak{a}_U(x)$ with respect to a dual basis in X_x. It is easily seen that the functions $\mathfrak{a}'_U$ define a C^∞ metric $\mathfrak{a}'$ for B'.

The tangent bundle T of $\mathfrak{V}$ is defined by taking for T_U the tangent bundle of U, for $h_U(g)$ the inverse of the mapping of tangent vectors induced by g, and for λ^* the inverse of the mapping of tangent vectors induced by the injection λ. If $\mathfrak{g}$ is a metric for T, then for each $\{U, G, \varphi\} \in \mathfrak{F}$ $\mathfrak{g}_U$ is a G-invariant Riemannian metric for U. The dual bundle T' of T is the bundle of differential 1-forms for $\mathfrak{V}$ which we henceforth denote by A^1. In general we use A^p to denote the bundle of differential p-forms for $\mathfrak{V}$, where, by definition, $A_U{}^p$ is the bundle of differential p-forms over U, $h_U(g) = g^*$, the mapping of differential forms dual to g, and λ^* is the mapping of differential forms dual to λ. The metric $\mathfrak{g}$ on T gives rise to a metric $\mathfrak{g}'_p$ in a well-known manner.

From now on we assume that the V-manifold $\mathfrak{V}$ is oriented. If $\{U, G, \varphi\}$ is a l. u. s., and ϕ is an n-form defined on an open subset $\mathcal{O}$ of $\varphi(U)$, we define the integral of ϕ over $\mathcal{O}$ by

$$\int_{\mathcal{O}} \phi = (\operatorname{ord} G)^{-1} \int_{\varphi^{-1}(\mathcal{O})} \phi_U.$$

From this it is obvious how we define the integral of an n-form over any measurable subset of $\mathfrak{V}$. It is easy to show that if η is any differentiable $(n-1)$-form with compact support on $\mathfrak{V}$, then $\int_{\mathfrak{V}} d\eta = 0$. We let $\{l_i\}$ be a C^∞ partition of unity subordinate to a locally finite covering $\{U_i\}$ of $\mathfrak{V}$ such that $\{U_i, G_i, \varphi_i\} \in \mathfrak{F}$. Then

$$\int_{\mathfrak{V}} d\eta = \int_{\mathfrak{V}} d(\sum_i l_i\eta) = \sum_i \int_{\mathfrak{V}} d(l_i\eta).$$

However, since the support of $l_i\eta$ is compact and contained in U_i,

$$\int_{\mathfrak{V}} d(l_i\eta) = (\operatorname{ord} G_i)^{-1} \int_{U_i} d(l_i'\eta) = 0$$

by application of the ordinary Stokes' theorem, where $l_i'(y) = l_i(\varphi_i(y))$, $y \in U_i$.

If in particular n is the real dimension of $\mathfrak{V}$, the n-form dV defined by $dV_U = (\mathfrak{g}'_{nU})^{-\frac{1}{2}} dx^1 \wedge \cdots \wedge dx^n$, where $x^1, \cdots, x^n$ are the coordinates in U, gives us a measure for $\mathfrak{V}$, which measure we also denote by dV. (Note that $\mathfrak{g}'_{nU}$ is a positive real-valued function; in fact if $\mathfrak{g}$ has the matricial form $(\mathfrak{g}_{ij})$ in the coordinates on U, $\mathfrak{g}'_{nU} = \det(\mathfrak{g}_{ij})^{-1}$.) Then if B is any vector V-bundle over $\mathfrak{V}$ with a given metric $\mathfrak{a}$ and ϕ, ψ are measurable sections of B, we define $\| \phi \|^2 = \int_{\mathfrak{V}} \mathfrak{a}(\phi, \phi) dV$, $(\phi, \psi) = \int_{\mathfrak{V}} \mathfrak{a}(\phi, \psi) dV$.

We now suppose $\mathfrak{V}$ to be supplied with a C^∞ Riemannian metrc $\mathfrak{g}$. If $\phi = \{\phi_U\}$ is a differentiable form of degree p on $\mathfrak{V}$, we take $d\phi$ to be defined by the collection $\{d\phi_U\}$ of differentiable forms of degree $(p+1)$, where $d\phi_U$ is the ordinary exterior derivative of ϕ_U; $d\phi$ is well-defined since d commutes with the mapping of differential forms dual to a mapping of one differentiable manifold into another. Moreover, for any p-form ϕ on $\mathfrak{V}$ we define $*\phi = \{*\phi_U\}$ to be the form of degree $(n-p)$ such that for any p-form ψ,

$$\psi_U \wedge *\phi_U = \mathfrak{g}'_p(\psi, \phi) dV_U,$$

where dV_U is the previously defined n-form representing the invariant measure on U. If ϕ is differentiable, we define $\delta\phi = (-1)^{np+n+1} * d * \phi$. Finally we define $\Delta = d\delta + \delta d$.

If ϕ and ψ are differentiable forms of degrees p and $p+1$ respectively, of which at least one has compact support on $\mathfrak{V}$, it follows from the identity

$$d(\phi \wedge *\psi) = d\phi \wedge *\psi - \phi \wedge *\delta\psi$$

that $(d\phi, \psi) = (\phi, \delta\psi)$. Hence if ϕ and ψ are twice differentiable p-forms with compact support, $(\Delta\phi, \psi) = (\phi, \Delta\psi)$.

Suppose now that $\mathfrak{V}$ is a complex analytic V-manifold of complex dimension n. For each $\{U, G, \varphi\} \in \mathfrak{F}$ the complexified tangent bundle $T_U{}^C$ of U splits into a direct sum $T_U{}^C = \mathfrak{T}_U \oplus \bar{\mathfrak{T}}_U$, where $\mathfrak{T}_U$ is a complex analytic bundle with C^n as fibre and where $\bar{\mathfrak{T}}_U$ is the bundle with C^n as fibre and whose coordinate transformations are the complex conjugates of those of $\mathfrak{T}_U$. This splitting is preserved under the complex analytic bundle maps $h_U(g)$, $g \in G$, and λ^* of T. Correspondingly, the bundle $A^p{}_U$, $p = 1, \cdots, n$, splits in an invariant way into a direct sum $\bigoplus_{r+s=p} A^{r,s}{}_U$ where $A^{r,s}{}_U$ is the bundle of forms of type (r, s) over U, and the exterior differentiation d also splits into $\partial + \bar{\partial}$ where both ∂ and $\bar{\partial}$ commute with all g^* and λ^*. A Hermitian metric ω on $\mathfrak{V}$ is simply a Hermitian metric for $\mathfrak{T} = \{\mathfrak{T}_U\}$, and gives rise to a metric for T, i.e., a Riemannian metric for $\mathfrak{V}$, in a canonical fashion; moreover, there is associated with ω a differential 2-form also denoted by ω,

and if $d\omega = 0$, we say that ω is Kähler. We denote by ω'_p the metric on A^p induced by ω. If $B = \{B_U\}$ is a V-bundle over $\mathfrak{V}$ with the complex plane C as fibre and non-zero complex numbers C^* as group, briefly, a complex line bundle, and if $\mathfrak{a}$ is a Hermitian metric for B, then $\mathfrak{a} \otimes \omega'_p = \{\mathfrak{a}_U \otimes \omega'_{pU}\}$ is a metric for $B \otimes A^p = \{B_U \otimes A^p{}_U\}$ whose bundle maps corresponding to elements g of G, $\{U, G, \varphi\} \in \mathcal{F}$, and injections λ are the tensor products of those of B and A^p. It is clear that $B \otimes A^p = \bigoplus_{r+s=p} B \otimes A^{r,s}$, the terms of the direct sum having an evident meaning. We denote by $\mathfrak{a} \otimes \omega'_{r,s}$ the metric $\mathfrak{a} \otimes \omega'_p$ restricted to $B \otimes A^{r,s}$. We also denote as before by dV_U the n-form defining the invariant measure on U associated with ω_U. A section of $B \otimes A^{r,s}$ is, by definition, a form of type (r, s) with coefficients in B. By $-B$ we mean the complex line bundle for which the non-zero holomorphic functions representing the coordinate transformations and bundle maps $h_U(g)$ and λ^* in particular coordinate systems are the reciprocals of those of B. Then if ϕ is a form of type (r, s) with coefficients in B, we let $\#\phi$ be the form of type $(n-r, n-s)$ with coefficients in $-B$ such that for every form ψ of type (r, s) with coefficients in B we have

$$\psi_U \wedge \#\phi_U = (\mathfrak{a} \otimes \omega'_{r,s})(\psi, \phi)\, dV_U$$

in U. If ϕ is differentiable, we take $\bar{\partial}\phi$ to be defined by the collection $\{\bar{\partial}\phi_U\}$ of forms of type $(r, s+1)$ with coefficients in B; it is easily seen that $\bar{\partial}\phi$ is well-defined, and in particular that the collection $\{\bar{\partial}\phi_U\}$ satisfies the conditions s(i) and s(ii). Moreover, we define $\mathfrak{D}\phi = -\#\bar{\partial}\#\phi$. Finally we put $\square = \mathfrak{D}\bar{\partial} + \bar{\partial}\mathfrak{D}$. It follows from the identity

$$d(\phi_U \wedge \#\psi_U) = \bar{\partial}\phi_U \wedge \#\psi_U - \phi_U \wedge \#\mathfrak{D}\psi_U$$

that if ϕ and ψ are differentiable forms with coefficients in B having compact support and of types (r, s) and $(r, s+1)$ respectively, then $(\phi, \mathfrak{D}\psi) = (\bar{\partial}\phi, \psi)$. Therefore, if ϕ and ψ are twice differentiable forms of type (r, s) with coefficients in B, then $(\square\phi, \psi) = (\phi, \square\psi)$.

3. Facts about geodesics. In this Section as in the following two Sections we let U be a small neighborhood of the origin ξ_0 in R^n and let G be a finite group of orientation-preserving linear transformations of U onto itself. Instead of the usual Euclidean metric we suppose given some G-invariant, C^∞ Riemannian metric in U. We let this metric be represented by the positive definite matrix function $(\mathfrak{g}_{ij})$ with respect to the coordinates $x^1, \cdots, x^n$ with center at ξ_0. Then Euler's equations for geodesic lines parametrized by $x^i = x^i(t)$ are

$$(3.1) \qquad d(L^{-\frac{1}{2}} \sum_k \mathfrak{g}_{jk} x^k{}_t)/dt - \tfrac{1}{2} \sum_{i,k} L^{-\frac{1}{2}} \partial_j \mathfrak{g}_{ik} x^i{}_t x^k{}_t = 0, \qquad j = 1, \cdots, n,$$

where $_t$ denotes differentiation with respect to t, ∂_j is partial differentiation with respect to x_j, and L is the Lagrangian, $L=\sum_{i,k} g_{ik}x^i{}_t x^k{}_t$. If we let s denote arclength, $s=\int_{t_0}^{t_1} L^{\frac{1}{2}}\,dt$, then the equations (3.1) become

$$(3.2)\qquad (d/ds)\big(\sum_k g_{jk}(dx^k/ds)\big)-\tfrac{1}{2}\sum_{i,k}\partial_j g_{ik}(dx^i/ds)(dx^k/ds)=0.$$

If we put $y_j=\sum_k g_{jk}(dx^k/ds)$, and $H=\frac{1}{2}\sum_{i,k} g^{ik}y_i y_k$, where $\sum_j g^{ji}g_{jk}=\delta^i{}_k$, we have

$$(3.3)\qquad x^k{}_t=(\partial H/\partial y_k),\qquad k=1,\cdots,n,$$

while our equations (3.2) become

$$(3.4)\qquad (y_k)_t=-\partial H/\partial x^k.$$

We then have the following easily verified facts: (See Kodaira [4] and de Rham [8], pp. 132-138).[2]

3(i) If $\{\xi^i\}$ and $\{\eta_i\}$, $i=1,\cdots,n$, are sufficiently small real numbers, there exists one and only one system of solutions $x^i=x^i(t;\xi,\eta)$ and $y_i=y_i(t;\xi,\eta)$ satisfying (3.3) and (3.4) and the initial conditions $x^i(0)=\xi^i$, $y_i(0)=\eta_i$; moreover, these solutions are C^∞ functions of ξ, η, and t.

3(ii) $x^i(t;\xi,\eta)=x^i(c^{-1}t;\xi,c\eta)$ and $y_i(t;\xi,\eta)=c^{-1}y_i(c^{-1}t;\xi,c\eta)$, so that if we put $x^i(\xi,\eta)=x^i(1;\xi,\eta)$ and $y_i(\xi,\eta)=y_i(1;\xi,\eta)$, we have

3(ii)′ $y_i(t;\xi,\eta)=t^{-1}y_i(\xi,t\eta)$ and $x^i(t;\xi,\eta)=x^i(\xi,t\eta)$.

3(iii) Letting $\eta^l=\sum_k g^{lk}\eta_k$ we have

$$x^i(\xi,\eta)=\xi^i+\eta^i+O(\eta^2),$$

or, what amounts to the same thing,

$$\eta^i=x^i-\xi^i+O(\eta^2)=x^i-\xi^i+O[(x-\xi)^2],$$

the terms $O(\ )$ and $O[\ ]$ denoting terms of the second order. The coordinates $\eta^1,\cdots,\eta^n$, having center ξ and valid in a small neighborhood of ξ, are called normal coordinates.

---

[2] The existence and uniqueness of solutions $x^i(t;\xi,\eta)$, $y_i(t;\xi,\eta)$ is assured by Theorem 2, p. 42 of [6]. That the solutions are of class C^∞ in ξ, η, t is assured by Theorem 6, p. 101 of [6]. The formula (ii) follows from (i) by direct calculation; (iii) follows from (ii) by Taylor's theorem with remainder in integral form. The rest of the formulas are a matter of direct calculation.

3(iv) Along a geodesic H is constant.

3(v) Letting $r(x,\xi)$ be the geodesic distance between x and ξ, and putting $P(x,\xi)=r(x,\xi)^2$ we have $P(x,\xi)=\sum\eta^i\eta_i=\sum y^iy_i$.

3(vi) $\partial P(x,\xi)/\partial x^i=2y_i$.

3(vii) $\sum P_iP^i=4P$, where $P_i=\partial P(x,\xi)/\partial x^i$ and

$$P^i(x,\xi)=\sum_k g^{ik}(x)P_k(x,\xi).$$

3(viii) if f is a differentiable function, the derivative $\partial f/\partial r$ of f along a geodesic is given by

$$2r\partial f/\partial r=\sum_i P^i(\partial f/\partial x^i),$$

r being geodesic distance.

3(ix) $$\sum_{k,m} g^{km}(x)\,(\partial^2P(x,\xi)/\partial x^k\partial x^m)=2n.$$

3(x) $$P_k(\xi,\xi)=0.$$

4. C^ν approximation of an elementary solution. We retain the notation of the preceding section and for each $g\in G$ we let $\mathfrak{M}_g=(\mathfrak{M}^I{}_{gJ})$, $I,J=1,\cdots,N$, be a C^∞ mapping of U into the manifold of (real or complex) non-singular $N\times N$ matrices such that if $x\in U$, $g,g'\in G$, then

$$\mathfrak{M}_{gg'}(x)=\mathfrak{M}_g(g'x)\mathfrak{M}_{g'}(x).$$

We let $\mathfrak{A}=(\mathfrak{A}^{IJ})$ be a fixed C^∞ mapping of U into the manifold of (symmetric or Hermitian) positive definite $N\times N$ matrices such that

$${}^t\mathfrak{M}_g(x)\mathfrak{A}(gx)\overline{\mathfrak{M}_g(x)}=\mathfrak{A}(x).$$

In other language, $g\to\mathfrak{M}_g$ defines an isomorphism of G into a group of covariant bundle maps of the bundle $U\times R^N$ (real case) or $U\times C^N$ (complex case) onto itself (by $(x,v)\to(gx,\mathfrak{M}_g(x)v)$), and $\mathfrak{A}$ is simply a G-invariant metric for the bundle. Let $\phi=(\phi_1,\cdots,\phi_N)$ be a (real or complex) vector-valued function on U, which we also view as an $N\times 1$ column matrix to simplify notation. We define the dual ϕ^* of ϕ to be the vector-valued function with components $(\phi^1,\cdots,\phi^N)$ defined by $\phi^I(x)=\sum_J\mathfrak{A}^{IJ}(x)\overline{\phi_J(x)}$, and if $\psi=(\psi_1,\cdots,\psi_N)$ is another N-dimensional vector-valued function, we define

$$(\phi,\psi)_U=\int_U {}^t\phi\cdot\psi^*dV=\int_U\sum_I\psi^I\phi_I\,dV,$$

where dV is the measure arising from the G-invariant metric and ${}^t\phi$ denotes

the transpose of ϕ, and define $\|\phi\|^2_U = (\phi, \phi)_U$, provided the integrals defining these quantities exist. In general ϕ_g is defined by $\phi_g(x) = \mathfrak{M}_g(x)^{-1}\phi(gx)$. We say that ϕ is G-invariant if $\phi_g(x) = \phi(x)$ for $g \in G$, $x \in U$. Then if ϕ and ψ are vector-valued functions, and if ϕ is G-invariant, we have

$$
\begin{aligned}
(4.1) \qquad {}^t\phi(x)\mathfrak{a}(x)\overline{\psi(x)} &= {}^t\phi(gx){}^t\mathfrak{M}_g(x)^{-1}\mathfrak{a}(x)\overline{\psi(x)} \\
&= {}^t\phi(gx)\mathfrak{a}(gx)\overline{\mathfrak{M}_g(x)\psi(x)} = {}^t\phi(gx)\mathfrak{a}(gx)\overline{\psi_g(gx)}
\end{aligned}
$$

and hence if ϕ and ψ are square summable in the sense that $\|\phi\|^2_U < +\infty$ and $\|\psi\|^2_U < +\infty$,

$$(\phi, \psi)_U = (\phi, \psi_g)_U, \quad g \in G,$$

since the measure dV is G-invariant.

We then suppose given a strongly elliptic differential operator ∇; which, by definition, means an operator which carries a vector-valued function ϕ of class C^2 into a vector-valued function $\nabla\phi$ defined by

$$(4.2) \qquad (\nabla\phi)_I = -\sum_{l,m} g^{lm}\partial_l\partial_m\phi_I + \sum_{m,J} a^{mJ}{}_I\partial_m\phi_J + \sum_J b^J{}_I\phi_J,$$

$a^{mJ}{}_I$ and $b^J{}_I$ being functions of class C^∞. We let A^m and B denote respectively the matrices $(a^{mJ}{}_I)$ and $(b^J{}_I)$. Then (4.2) becomes

$$(4.3) \qquad \nabla\phi = -\sum g^{lm}\partial_l\partial_m\phi + \sum_m A^m\partial_m\phi + B\phi.$$

We note here that all our dealing for the time being have only to do with a fixed system of coordinates and that $a^{mJ}{}_I$ and $b^J{}_I$ are not generally invariantly defined. We remark in passing that the Laplacian $\Delta = \delta d + d\delta$ is strongly elliptic on the module of differential forms of degree k. See [4].

In what follows we use $\phi \equiv \psi \bmod \mathfrak{N}$ to mean "ϕ equals ψ almost everywhere," and $\phi \in A \bmod \mathfrak{N}$ means there is a member of the class of functions A to which ϕ is equal almost everywhere. Moreover, if $K(x, \xi)$ is an $N \times N$ matrix function on $U \times U$, $\nabla_x K(x, \xi)$ is the $N \times N$ matrix function whose J-th column is obtained by applying ∇ to the J-th column of $K(\ , \xi)$ for a fixed value of ξ. $\int K(x, \xi)dV_x$ means that we integrate each entry of K with respect to x holding ξ fixed.

Our first problem in this Section is to find an $N \times N$ matrix function $\bar{\mathbf{\Gamma}}^\nu = \{\Gamma^{\nu J}{}_I\}$ such that:

(a) $\nabla_x \bar{\mathbf{\Gamma}}^\nu(x, \xi)$ is a matrix function whose entries are real functions which belong to $C^\nu \bmod \mathfrak{N}$ on $U \times U$.

$$\text{(b)}\quad \tilde{\mathbf{\Gamma}}^{\nu}(x,\xi)=\begin{cases}[(n-2)\omega_n]^{-1}P(x,\xi)^{-\tau}U(x,\xi)+\log P(x,\xi)V(x,\xi), & n\geqq 3,\\ -1/4\pi\log P(x,\xi)U(x,\xi)+V(x,\xi),\quad n=2, & x\neq\xi,\end{cases}$$

where $\tau=(n-2)/2$, where ω_n is the "area" of the unit sphere in R^n, U and V are matrix functions of class C^∞ in x and ξ, and $U(\xi,\xi)$ is the identity matrix which we denote by $\mathbf{1}$. To achieve this end we tentatively seek to replace the condition (a) by the stronger condition

$$\text{(a)}'\qquad \nabla_x\tilde{\mathbf{\Gamma}}(x,\xi)=0$$

and begin by writing $\tilde{\mathbf{\Gamma}}=\{\tilde{\mathbf{\Gamma}}^{\nu}{}_{I}(x,\xi)\}$ so that we have

$$(4.4)\qquad \nabla_x\tilde{\mathbf{\Gamma}}=-\sum g^{lm}\partial_l\partial_m\tilde{\mathbf{\Gamma}}+\sum_m A_m\cdot\partial_m\tilde{\mathbf{\Gamma}}+B\cdot\tilde{\mathbf{\Gamma}},$$

all terms being $N\times N$ matrices and multiplication being matrix multiplication. We then seek to find U and V in the form

$$\text{(c)}\qquad \begin{cases}U=M\{\mathbf{1}+\sum_{\kappa=1}^{\infty}P^{\kappa}U_{\kappa}\},\\ V=M\{\sum_{\kappa=0}^{\infty}P^{\kappa}V_{\kappa}\}, \qquad P^{\kappa}=[P(x,\xi)]^{\kappa},\end{cases}$$

where M is a matrix which we will shortly define and where U_κ and V_κ are matrix functions of class C^∞ which we will find by a simple recursion formula.

Let $l(P)$ be any (twice differential) scalar function of $P=P(x,\xi)$ and let W be a (twice differentiable) $N\times N$ matrix function of x and ξ. Then direct computation gives

$$(4.5)\quad \nabla_x(l(P)\cdot W)=l(P)\nabla_x W+l'(P)\{-\sum g^{km}\partial_k\partial_m P+\sum A^kP_k-4r\frac{\partial}{\partial r}\}W \\ -l''(P)4P\cdot W.$$

We set in turn

$$(4.6)\qquad \sum g^{km}\partial_k\partial_m P-\sum A^kP_k=2n-4\mathfrak{R}(x,\xi),$$

$$(4.7)\qquad \mathfrak{R}^*(\xi,\eta)=\mathfrak{R}(x(\xi,\eta),\xi),$$

$$(4.8)\qquad L(\xi,\eta,t)=t^{-1}\mathfrak{R}^*(\xi,t\eta),$$

and obtain immediately from 3(ix) and 3(x) above, that $\mathfrak{R}(\xi,\xi)=0$, $\mathfrak{R}^*(\xi,0)=0$, from which it follows[3] that $L(\xi,\eta,t)$ is a function of t,ξ,η of class C^∞; moreover, (4.5) becomes

$$(4.9)\quad \nabla_x(l(P)\cdot W)=l(P)\cdot\nabla_x W-l'(P)\{2n-4\mathfrak{R}(x,\xi)+4r\frac{\partial}{\partial r}\}W \\ -4l''(P)\cdot P\cdot W.$$

[3] This is a consequence of Taylor's theorem with remainder in integral form.

Now put

$$M(x,\xi)=M(x(\xi,t\eta),\xi)=\mathbf{1}+\int_0^t L(t_1)\,dt_1+\int_0^t\int_0^{t_1}L(t_2)L(t_1)\,dt_2dt_1$$
$$+\cdots+\int_0^t\int_0^{t_1}\cdots\int_0^{t_{m-1}}L(t_m)\cdots L(t_1)\,dt_m\cdots dt_1+\cdots,$$

where $L(t_i)=L(\xi,\eta,t_i)$. That M is actually a function of x and ξ follows from 3(ii), 3(ii)′, and 3(iii), and that the series converges can be seen in the following way: If $|L|$ denotes the norm of the matrix L (see [4], p. 613) and if $|L(t,\xi,\eta)|<K\cdot r(x,\xi)$ for $0\leqq t\leqq 1$, $|\xi|<\epsilon$, $|\eta|<\epsilon'$, the $(m+1)$-th term of this series has norm less than $m!^{-1}(K^m r^m)$ in this region. Thus the series converges absolutely in this region, and since the series when differentiated term by term with respect to t also converges uniformly, it follows that $dM/dt=LM=t^{-1}\mathfrak{R}^*(\xi,\eta t)M$, or along a geodesic $r(\partial M/\partial r)=\mathfrak{R}(x,\xi)M$. Therefore, if we set $\nabla^*=M^{-1}\nabla_x M$, (4.8) becomes

$$M^{-1}\nabla_x(l(P)MW)=l\nabla^*W-l'(P)\{2n+4r\frac{\partial}{\partial r}\}W-4Pl''(P)W. \tag{4.10}$$

We refer to equations (b) and (c) and consider 3 cases. If n is odd, we put all $V_\kappa=0$, for notational convenience put $U_{-1}=0$, and obtain

$$-M^{-1}\nabla_x\Gamma=\sum_{\kappa=0}^{\infty}P^{\kappa-\tau+1}\{4(\kappa-\tau)[\kappa+r\frac{\partial}{\partial r}]U_\kappa-\nabla^*U_{\kappa-1}\};$$

then

$$4(\kappa-\tau)[\kappa+r\frac{\partial}{\partial r}]U_\kappa-\nabla^*U_{\kappa-1}=0 \tag{4.11}$$

is satisfied if we take $U_0=\mathbf{1}$ and define U_κ for $\kappa>0$ by

$$U_\kappa(x,\xi)=[4(\kappa-\tau)r(x,\xi)^\kappa]^{-1}\int_0^{r(x,\xi)}s^{\kappa-1}\nabla^*U_{\kappa-1}ds,$$

i.e., if $x=x(\xi,\eta)$ and if we put $t=s/r$,

$$U_\kappa(x,\xi)=[4(\kappa-\tau)]^{-1}\int_0^1 t^{\kappa-1}\nabla^*_x U_{\kappa-1}(x(\xi,t\eta),\xi)\,dt$$

from which it is evident that all U_κ are of class C^∞. If n is even and >2, U_κ and V_κ are determined by similar recursion formulae (see [4], p. 614) in such a manner that the partial sums in the corresponding expansion of $M^{-1}\nabla_x\Gamma$ are made to "telescope." Finally, the case $n=2$ is to be treated separately, and it too gives rise to similar recursion formulae (see [4], p. 615). In each case we choose $U_0=\mathbf{1}$. Then if n is odd we put

$$\tilde{\Gamma}^\nu(x,\xi)=P^{-\tau}M\sum_{\kappa=0}^{\nu+\frac{1}{2}(n-1)}P^\kappa U_\kappa,$$

from which we obtain, using (4.11),

$$\nabla_x \tilde{\mathbf{\Gamma}}^\nu = MP^\nu \sqrt{P}\nabla^*{}_x U_{\nu+\frac{1}{2}(n-1)}$$

which is of class C^ν. If n is even we put

$$\tilde{\mathbf{\Gamma}}^\nu = P^{-\tau} M\{\mathbf{1} + \sum_{\kappa=1}^{\nu+1+\tau} P^\kappa U_\kappa\} + M\cdot\log P \sum_{\kappa=0}^{\nu+1} P^\kappa V_\kappa$$

which gives

$$\nabla_x \tilde{\mathbf{\Gamma}}^\nu = M\{P^{\nu+1}\nabla^*{}_x U_{\tau+\nu+1} + (\log P)\cdot P^{\nu+1}\nabla^*{}_x V_{\nu+1}\},$$

which is of class C^ν. Thus we have found $\tilde{\mathbf{\Gamma}}^\nu$ satisfying (a) and (b). It is at this point that the procedure of [4] is modified to eliminate discussion of the convergence of the infinite series appearing in (c) which is established in [4] only for the analytic case; and it is our desire to discuss the C^∞ case. $\tilde{\mathbf{\Gamma}}^\nu$ has been called by Kodaira a C^ν approximation to an elementary solution.

For suitably small $\epsilon > 0$ we let U_1 be a neighborhood of ξ_0 such that if $x \in U_1$ and $r(x,\xi)^2 < 3\epsilon$, then $\xi \in U$. We let ρ be a non-negative, real-valued function on the positive reals such that $0 \leqq \rho(y) \leqq 1$ for all y and

$$\rho(y) = 1 \text{ if } y \leqq \epsilon, \text{ and } \rho(y) = 0 \text{ if } y \geqq 2\epsilon,$$

and then define

$$\Gamma_*{}^\nu(x,\xi) = \rho(r(x,\xi)^2)\tilde{\mathbf{\Gamma}}^\nu(x,\xi).$$

From now on we assume that ∇ commutes with $\mathfrak{M}_g$, i.e.,

$$\mathfrak{M}_g(x)(\nabla\phi)(x) = (\nabla\phi_{g^{-1}})(gx),$$

for each $g \in G$ and every (twice differentiable) ϕ.

We then further define

$$\Gamma^\nu(x,\xi) = (\operatorname{ord} G)^{-1}\sum_{g\in G}\mathfrak{M}_g(x)^{-1}\Gamma_*{}^\nu(gx,g\xi)\,\mathfrak{M}_g(\xi), \quad x \neq \xi,$$

so that $\mathfrak{M}_g(x)\Gamma^\nu(x,\xi) = \Gamma^\nu(gx,g\xi)\mathfrak{M}_g(\xi)$. Then $Q^\nu(x,\xi)$, is defined by $Q^\nu(x,\xi) = \nabla_x\Gamma^\nu(x,\xi)$. It is clear from the construction of Γ^ν that Q^ν is of class C^ν. Since

$$\Gamma_*{}^\nu(x,\xi) = [(n-2)\omega_n r(x,\xi)^{n-2}]^{-1}\{M(x,\xi) + O(r(x,\xi))\},$$

where $M(\xi,\xi) = \mathbf{1}$, and since the metric and therefore $r(x,\xi)$ are G-invariant, we have

(4.12) $\quad \Gamma^\nu(x,\xi) = [(n-2)\omega_n r(x,\xi)^{n-2}]^{-1}\{M^*(x,\xi) + O(r(x,\xi))\},$

where $M^*(\xi,\xi)$ is the identity. The case $n = 2$ is slightly different in form,

but we will not give it special attention except when necessary. It should be pointed out here that $\Gamma^\nu(x,\xi)$ and $Q^\nu(x,\xi)$ are not symmetrical in x and ξ.

5. The regularity theorem and auxiliary lemmas. Retaining the notation of the previous sections, we define $\mathfrak{H}$ as the module of vector-valued functions ϕ on U such that $\|\phi\|_U < +\infty$ and $\mathfrak{H}[U_1]$ as the module of vector-valued functions ϕ on U_1 such that $\|\phi\|_U < +\infty$, while we let

$$\mathfrak{H}_\nu = \{\phi \mid \phi \in \mathfrak{H}, \phi \in C^\nu\} \quad \text{and} \quad \mathfrak{H}_\nu[U_1] = \{\phi \mid \phi \in \mathfrak{H}[U_1], \phi \in C^\nu\}.$$

Then $\mathfrak{H}$ and $\mathfrak{H}[U_1]$ with the norms $\|\ \|_U$ and $\|\ \|_{U_1}$ give rise to Hilbert spaces which we also denote by $\mathfrak{H}$ and $\mathfrak{H}[U_1]$, but in which we identify two functions which are equal almost everywhere. It is trivial that the subspace of G-invariant functions are closed since the group G acts continuously in $\mathfrak{H}$ and $\mathfrak{H}[U_1]$.

In what follows it would be possible to use certain formulas from [8], pp. 132-159, if we were speaking of differential forms only. Inasmuch as it is more convenient for our purposes to speak of vector-valued functions instead of differential forms, de Rham's formulas would have to be modified slightly to apply here. The point is that we wish to include the case of differential forms with coefficients in a vector bundle, and in order to avoid complicated indicial notation it seemed desirable to us to speak of vector-valued functions. As remarked in [8], p. 35, however, no really essential changes are needed to pass over to the case of vector-valued forms, and the reader who desires to do so may interpret what we say here in the notation of de Rhams's book without great difficulty.

We define the integral operators Γ^ν and Q^ν by

$$\overline{(\Gamma^\nu\phi)(\xi)} = \mathfrak{A}(\xi)^{-1}\int_U {}^t\Gamma^\nu(x,\xi)\,\mathfrak{A}(x)\overline{\phi(x)}\,dV_x$$

and

$$\overline{(Q^\nu\phi)(\xi)} = \mathfrak{A}(\xi)^{-1}\int_U {}^tQ^\nu(x,\xi)\,\mathfrak{A}(x)\overline{\phi(x)}\,dV_x$$

provided that the integrands are summable.

Using this notation we wish to prove the following theorems and lemmas from which our desired results follow easily:

LEMMA α. *Suppose μ is a measurable real function on the non-negative real axis such that $|\mu(r)| \leqq Cr^{1-n}$, $C > 0$, and such that $\mu(r) = 0$ if $r^2 \geqq 2\varepsilon$.*

Let $W(x,\xi)=\{W(x,\xi)^I{}_J\}$ *be a matrix function of class* C^∞ *on* $U\times U_1$. *Suppose* $\phi\in C^\infty$, $\mathfrak{L}(\phi)\subset U_1$, *where* $\mathfrak{L}$ *denotes support. Then* ψ, *defined by*

$$\psi(\xi)=\int_U \mu(r(x,\xi))W(x,\xi)\phi(x)\,dV_x,\quad \xi\in U,$$

is of class C^∞ *and* $\mathfrak{L}(\psi)\subset U$.

(The most essential part of Lemma α, namely that ψ is of class C^∞, can be obtained under the stronger hypothesis that μ is $0(r^{2-n})$, and this is sufficient for our purposes in view of (4.12). This weaker result is practically the same as part of Lemma 4, p. 138 of [8].)

LEMMA β. *Suppose* β *is* C^∞, $\mathfrak{L}(\beta)\subset U_1$. *Then* $\int_U \Gamma^\nu(x,\xi)\beta(\xi)\,dV_\xi$ *is of class* C^∞ *in* x.

LEMMA γ. *If* $\phi\in\mathfrak{J}$, $Q^\nu\phi\in\mathfrak{J}_\nu[U_1]$.

LEMMA δ. Q^ν *is a completely continuous mapping of* $\mathfrak{J}$ *into* $\mathfrak{J}_\nu[U_1]$.

LEMMA ϵ. Γ^ν *is a bounded mapping of* $\mathfrak{J}_\infty$ *into* $\mathfrak{J}_\infty[U_1]$.

Before stating two more theorems, we remark that ∇ has an adjoint ∇^* such that for all (twice differentiable) $\phi,\psi\in\mathfrak{J}$, $(\phi,\nabla\psi)=(\nabla^*\phi,\psi)$, provided either ϕ or ψ has compact support contained in U. In fact, ∇^* is defined by

$$\begin{aligned}\nabla^*\phi &= (\mathfrak{g})^{\frac12}\mathfrak{A}^{-1}[-\sum_{l,m}\partial_l\partial_m(\mathfrak{g}^{lm}(\mathfrak{g})^{\frac12}\mathfrak{A}\phi)-\sum_m\partial_m(A^m(\mathfrak{g})^{\frac12}\mathfrak{A}\phi)+(\mathfrak{g})^{\frac12}B\mathfrak{A}\phi]\\ &= -\sum_{l,m}\mathfrak{g}^{lm}\partial_l\partial_m\phi+\sum_m A^m{}_*\partial_m\phi+B_*\phi,\end{aligned}$$

where $\mathfrak{g}=\det(\mathfrak{g}_{ij})$ is the determinant of the Riemannian metric form in the given system of coordinates. This formula may be verified by repeated integration by parts. The main fact we need is that the sum of the terms involving second partial derivatives of ϕ is $-\sum_{l,m}\mathfrak{g}^{lm}\partial_l\partial_m\phi$, and since we already know that $\Delta^*=\Delta$ and $\Box^*=\Box$ (see Section 2), we do not stop to verify the above formula here.

THEOREM A_N. *Let* N *have the same meaning as at the beginning of Section* 4. *Let* $\phi\in\mathfrak{J}$ *be such that* $\mathfrak{L}(\phi)\subset U_1$. *Put*

$$\psi(x)=\int_U \Gamma^\nu(x,\xi)\phi(\xi)\,dV_\xi.$$

Then $\psi\in C^\infty$, $\mathfrak{L}(\psi)\subset U$, and

$$\nabla\psi(x)=\phi(x)+\int_U Q^\nu(x,\xi)\phi(\xi)\,dV_\xi,\quad x\in U.$$

THEOREM A^*_N. *Let $\eta \in \mathfrak{J}_\infty$ be such that $\mathfrak{L}(\eta) \subset U$. Then*

$$\Gamma^\nu \nabla^* \eta(\xi) = \eta(\xi) + Q^\nu \eta(\xi), \quad \xi \in U_1.$$

The proofs of Lemmas α-ϵ are relatively simple. As for the proof of Theorems A_N and A^*_N, we shall show that if we assume Lemma α, Theorem A_N is equivalent to Theorem A^*_N; we shall then show that Thecrem A_0 implies Theorem A_N, and shall finally prove Theorem A_0^*.

Let U and S be open subsets of Euclidean space and let $F(x,y)$ be a measurable matrix function on $S \times U$. We say that $\int_U F(x,y)\,dV_y$ converges boundedly on S if $\int_U |F(x,y)|\,dV_y < M$ for all $x \in S$. We will need the following preparatory

LEMMA 0. *Let $F(x,y)$ be a matrix function on $S \times U$ and continuously differentiable in x for $x \neq y$ such that $\partial F/\partial x^i$ is absolutely summable over $K \times U$ for any compact 1-dimensional interval K in S, $i = 1, \cdots, n$. Then $\partial[\int_U F(x,y)\,dV_y]/\partial x^i$ exists and equals $\int_U [\partial F(x,y)/\partial x^i]\,dV_y$ gor $x \in S$. In particular, this is true if $\int_U [\partial F(x,y)/\partial x^i]\,dV_y$ converges boundedly on S.*

Proof. Let $a = (a^1, \cdots, a^n) \in S$ and let $x = (a^1, \cdots, x^i, \cdots, a^n) \in S$ be such that the straight line segment $[a,x]$ lies in S. Then for fixed y

$$\int_a^x [\partial F(u,y)/\partial u^i]\,du^i = F(x,y) - F(a,y) \ \text{ if } \ y \notin [a,x].$$

Since $[a,x]$ is a set of measure zero,

$$\int_U \int_a^x [\partial F(u,y)/\partial u^i]\,du^i dV_y = \int_U [F(x,y) - F(a,y)]\,dV_y$$
$$= \int_U F(x,y)\,dV_y + \text{const.}$$

Then by Fubini's theorem,

$$\int_a^x \int_U [\partial F(u,y)/\partial u^i]\,dV_y\,du^i = \int_U F(x,y)\,dV_y + \text{const.}$$

Taking the derivative of both sides, we have

$$\int_U [\partial F(x,y)/\partial x^i]\,dV_y = \partial[\int_U F(x,y)\,dV_y]/\partial x^i. \qquad \text{q.e.d.}$$

1. *Proof of Lemma α.* By assuming U suitably small we may assume

that there exist C^∞ functions $h^\nu{}_j$, $\nu, j = 1, \cdots, n$ in U such that for each $\xi \in U$ we have $g_{jk}(\xi) = \sum_{\nu=1}^{n} h^\nu{}_j(\xi) h^\nu{}_k(\xi)$. This is true since $\{g_{jk}\}$ is positive definite and of class C^∞. We then let $\eta^1, \cdots, \eta^n$ be normal geodesic coordinates with center ξ and write $\theta^\nu(x, \xi) = \sum h^\nu{}_j(\xi)\eta^j(x, \xi)$. Since $\{h^\nu{}_j(\xi)\}$ is a non-singular C^∞ matrix, $\theta^1, \cdots, \theta^n$ can be used as local coordinates. (Such coordinates are called orthonormal coordinates.) Then

$$r(x, \xi)^2 = \sum_{\nu=1}^{n} (\theta^\nu(x, \xi))^2.$$

Moreover, $x^i = x^i(\xi, \theta)$ is a C^∞ function of ξ and θ. Therefore

$$\psi(\xi) = \int_U \mu(r(x, \xi)) W(x, \phi(x) dV_x$$
$$= \int_U X(\xi, \theta)\mu([\sum(\theta^\nu)^2]^{\frac{1}{2}}) d\theta^1 \cdots d\theta^n,$$

where X is a C^∞ matrix function of ξ and θ. Since $\mu(r) = 0$ if $r^2 > 2\epsilon$, and since $|\mu(r)| \leqq K r^{1-n}$

$$\int_{r^2 \leqq 2\epsilon} \mu([\sum(\theta^\nu)^2]^{\frac{1}{2}}) d\theta^1 \cdots d\theta^n \leqq \int_{r^2 \leqq 2\epsilon} K r^{1-n} d\theta^1 \cdots d\theta^n < M < +\infty.$$

Therefore, not only does the integral expressing $\psi(\xi)$ converge absolutely and boundedly for $\xi \in U$, since the support of ϕ is contained in U_1, but so also do the integrals $\int_{r^2 \leqq 2\epsilon} Y(\xi, \theta)\mu([\sum \theta^\nu)^2]^{\frac{1}{2}}) d\theta^1 \cdots d\theta^n$, where $Y(\xi, \theta)$ is any of the mixed partial derivatives of $X(\xi, \theta)$ with respect to $\xi^1, \cdots, \xi^n$. Hence, by using Lemma 0, it is easily seen that integration and differentiation may be interchanged, and therefore ψ possesses continuous derivatives of all orders.

2. *Proof of Lemma β.* This is an immediate corollary to Lemma α since $\Gamma^\nu(x, \xi)$ is a sum of a finite number of terms having the form $\mu(r(x, \xi)) W(x, \xi)$ considered in Lemma α.

3. *Proof of Lemma γ.* We only need show that the integral

$$\vartheta(\xi) = \int_U {}^tQ(x, \xi)\mathfrak{A}(x)\overline{\phi(x)} dV_x$$

is differentiable of class C^ν in U_1 provided ϕ is square summable over U. However, $Q^\nu(x, \xi) = 0$ if $r(x, \xi)^2 \geqq 2\epsilon$, while if $\xi \in U_1$ and $r(x, \xi)^2 \leqq 3\epsilon$, then $x \in U$. Hence, if Y is one of the mixed partial derivatives of $Q^\nu(x, \xi)$

with respect to $\xi^1, \cdots, \xi^n$ of order $\leqq \nu$, then since Y is continuous in both variables, the integral

$$\int_U Y(x,\xi)\mathfrak{a}(x)\overline{\phi(x)}dV_x$$

converges absolutely, and boundedly, as we see using Schwarz's inequality, for $x \in U_1$. Therefore, integration and differentiation are interchangeable, and therefore $\mathfrak{J} \in C^\nu$. q. e. d.

4. *Proof of Lemma δ.* This lemma follows immediately from the facts that the kernel $Q^\nu(x,\xi)\mathfrak{a}(x)$ of Q^ν is uniformly continuous over $(U_1 + 2\epsilon) \times U_1$, where $U_1 + 2\epsilon = \{x \mid \liminf_{\xi \in U_1} r(x,\xi)^2 \leqq 2\epsilon\}$ (in fact, Q^ν is continuous on $U \times U$), and that $Q(x,\xi) = 0$ if $r(x,\xi)^2 \geqq 2\epsilon$. See Banach [2], p. 97.

5. *Proof of Lemma ε.* We wish to prove that there exists $K > 0$ (depending on ν) such that if $\phi \in \mathfrak{J}_\infty$, then $\| \Gamma^\nu \phi \|_{U_1} \leqq K \| \phi \|_U$. (See Kodaira [4], p. 624.) If P is a continuous positive definite (symmetric or Hermitian) $N \times N$ matrix function on U, we define

$$M(P(x)) = \underset{\Sigma_I |z_I| = 1}{\text{maximum}} \, (\textstyle\sum P_{IJ}(x) z_I \bar{z}_J), \quad x \in U,$$

and let $M(P)$ be the least upper bound of the numbers $M(P(x))$ for $x \in U_1 + 2\epsilon$. We let ${}^*\Gamma^\nu(x,\xi) = \Gamma^\nu(x,\xi)\mathfrak{a}(x)$ and define

$$| {}^*\Gamma^\nu(x,\xi) | = \sqrt{\textstyle\sum_{I,J} | {}^*\Gamma^\nu(x,\xi)^{IJ} |^2}.$$

Since $| {}^*\Gamma^\nu(x,\xi) | \leqq K r^{1-n}$ uniformly on $U_1 + 2\epsilon$ and since $\Gamma^\nu(x,\xi) = 0$ if $r(x,\xi)^2 \geqq 2\epsilon$, the functions

$$f(\xi) = \int_U | {}^*\Gamma^\nu(x,\xi) | dV_x \quad \text{and} \quad g(x) = \int_{U_1} | {}^*\Gamma^\nu(x,\xi) | dV_\xi$$

are respectively bounded on U_1 and U. We let K_1 be a common bound for f in U_1 and g in U. Then by repeated application of Schwarz's inequality and by Fubini's theorem we have:

$$\begin{aligned}
\| \Gamma^\nu \phi \|^2_{U_1} &= \int_{U_1} {}^t(\Gamma^\nu \phi(\xi))\mathfrak{a}(\xi)(\overline{\Gamma^\nu \phi(\xi)})dV_\xi \\
&= \int_{U_1} {}^t\Big(\int_U \Gamma^\nu(x,\xi)\mathfrak{a}(x)\overline{\phi(x)}dV_x\Big)\mathfrak{a}(\xi)^{-1}\Big(\int_U \overline{\Gamma^\nu(x,\xi)\mathfrak{a}(x)}\phi(x)dV_x\Big)dV_\xi \\
&\leqq M(\mathfrak{a}^{-1}) \int_{U_1} {}^t\Big(\int_U \Gamma^\nu(x,\xi)\mathfrak{a}(x)\overline{\phi(x)}dV_x\Big)\Big(\int_U \overline{\Gamma^\nu(x,\xi)\mathfrak{a}(x)}\phi(x)dV_x\Big)dV_\xi \\
&\leqq M(\mathfrak{a}^{-1}) \int_{U_1} \Big(\int_U | {}^*\Gamma^\nu(x,\xi) | \, dV_x \cdot \int_U | {}^*\Gamma^\nu(x,\xi) | \, | \phi(x) |^2 dV_x\Big)dV_\xi
\end{aligned}$$

$$\leqq K_1 \cdot M(\mathfrak{A}^{-1}) \int_U \left(\int_{U_1} | {}^*\Gamma^\nu(x,\xi) | \, dV_\xi \right) | \phi(x) |^2 \, dV_x$$

$$\leqq NK_1^2(\mathfrak{A}^{-1}) \int_U | \phi(x) |^2 \, dV_x$$

$$\leqq K_1^2 M(\mathfrak{A}^{-1}) \frac{1}{m(\mathfrak{A})} \int_U {}^t\phi(x) \mathfrak{A}(x) \overline{\phi(x)} dV_x = K_3 \| \phi \|^2_U,$$

where $m(\mathfrak{A}) = \underset{x \in U+2\epsilon}{\text{g. l. b.}} (\underset{\Sigma_I |z_I|^2 = 1}{\text{minimum}} \sum \mathfrak{A}(x)^{IJ} z_I \bar{z}_J)$, which is our desired result.

6. *Proof that Theorem* A^*_N *is equivalent to Theorem* A_N. Suppose that $K(x,\xi)$ is an absolutely summable (entry by entry) matrix function on $(U_1 + 2\epsilon) \times U_1$ and that $K(x,\xi) = 0$ if $r(x,\xi)^2 \geqq 2\epsilon$. Let $\eta, \phi \in \mathfrak{J}_\infty$ and suppose that $\mathfrak{L}(\phi) \subset U_1$. Then by Fubini,

$$\int_{U_1} {}^t\phi(\xi) \left(\int_U K(x,\xi) \mathfrak{A}(x) \eta(x) dV_x \right) dV_\xi = \int_U {}^t\left(\int_{U_1} {}^tK(x,\xi) \phi(\xi) dV_\xi \right) \mathfrak{A}(x) \eta(x) dV_x.$$

Hence we have

$$\int_{U_1} {}^t\phi(\xi) \left(\int_U {}^t\Gamma^\nu(x,\xi) \mathfrak{A}(x) \overline{\nabla^* \eta(x)} dV_x \right) dV_\xi$$
$$= \int_U {}^t\{ \nabla_x \int_{U_1} \Gamma^\nu(x,\xi) \phi(\xi) dV_\xi \} \mathfrak{A}(x) \overline{\eta(x)} dV_x,$$

$$\int_{U_1} {}^t\phi(\xi) \left(\int_U {}^tQ^\nu(x,\xi) \mathfrak{A}(x) \overline{\eta(x)} dV_x \right) dV_\xi$$
$$= \int_U {}^t\{ \int_{U_1} Q^\nu(x,\xi) \phi(\xi) dV_\xi \} \mathfrak{A}(x) \overline{\eta(x)} dV_x,$$

and

$$\int_{U_1} {}^t\phi(\xi) (\mathfrak{A}(\xi) \overline{\eta(\xi)}) dV_\xi = \int_U {}^t\{ \phi(x) \} \mathfrak{A}(x) \overline{\eta(x)} dV_x.$$

Therefore

$$\int_{U_1} {}^t\phi(\xi) \left[\int_U {}^t\Gamma^\nu(x,\xi) \mathfrak{A}(x) \overline{\nabla^* \eta(x)} dV_x - \int_U {}^tQ^\nu(x,\xi) \mathfrak{A}(x) \overline{\eta(x)} dV_x - \mathfrak{A}(\xi) \overline{\eta(\xi)} \right] dV$$
$$= \int_U {}^t\{ \nabla_x \int_{U_1} \Gamma^\nu(x,\xi) \phi(\xi) dV_\xi - \int_{U_1} Q^\nu(x,\xi) \phi(\xi) dV_\xi - \phi(x) \} \mathfrak{A}(x) \overline{\eta(x)} dV_x.$$

Because of Lemmas α and δ we see that the right hand side of this equation vanishes for all $\eta, \phi \in \mathfrak{J}_\infty$ such that $\mathfrak{L}(\phi) \subset U_1$ if and only if Theorem A_N is true. while the left hand side vanishes for all such η and ϕ if and only if Theorem A^*_N is true. Hence the equivalence of the two theorems is established.

7. *Proof that Theorem* A_0 *implies Theorem* A_N. Let $K(x,\xi)$ be an $N \times N$ matrix function in U which is differentiable in x. Let $\phi \in \mathfrak{J}_\infty$, $\mathfrak{L}(\phi) \subset U_1$. Suppose that $\int_U [\partial K(x,\xi)/\partial x^k] \phi(\xi) dV_\xi$ converges absolutely,

and boundedly for x in any compact subset of U_1. Then from Lemma 0 we obtain

$$\partial\left[\int_U K(x,\xi)\phi(\xi)dV_\xi\right]/\partial x^k = \int_U [\partial K(x,\xi)/\partial x^k]\phi(\xi)dV_\xi, \text{ for } x\in U_1.$$

By hypothesis

$$(5.1) \qquad (\nabla\phi)_I = -\sum_{l,m} g^{lm}\partial_l\partial_m\phi_I + \sum_{m,J} a^{mJ}{}_I\partial_m\phi_J + \sum_J b^J{}_I\phi_J$$

and by construction

$$\Gamma^\nu(x,\xi) = \begin{cases} [(n-2)\omega_n r(x,\xi)^{n-2}]^{-1}\{\mathbf{1}+E_n(x,\xi)\} \\ \qquad = [(n-2)\omega_n r(x,\xi)^{n-2}]^{-1}\cdot\mathbf{1} + A_n(x,\xi), \quad n\neq 2. \\ (4\pi)^{-1}\log r(x,\xi)\{\mathbf{1}+E_2(x,\xi)\} \\ \qquad = (4\pi)^{-1}\log r(x,\xi)\cdot\mathbf{1} + A_2(x,\xi), \quad n=2. \end{cases}$$

The point we wish to make is that

$$|\partial A_n(x,\xi)/\partial x^k| \leqq C_1 r(x,\xi)^{1-n}, \qquad |\partial^2 A_n(x,\xi)/\partial x^k\partial x^j| \leqq C_2 r(x,\xi)^{1-n},$$

and since $\Gamma^\nu(x,\xi))=0$ if $r(x,\xi)^2\geqq 2\epsilon$, we see that $\partial_k A_n(x,\xi)$ and $\partial_k\partial_j A_n(x,\xi)$ are absolutely summable over U with respect to ξ for $x\in U_1+\epsilon$, and boundedly for $x\in U_1$. Moreover, $|\Gamma^\nu(x,\xi)|\leqq C_3 r(x,\xi)^{1-n}$ and $|\partial_k\Gamma^\nu(x,\xi)|\leqq C_4 r(x,\xi)^{1-n}$. Hence by our previous remarks, the J-th component of

$$\nabla_x\left[\int_{U_1}\Gamma^\nu(x,\xi)\phi(\xi)dV_\xi\right] - \int_{U_1}\nabla_x\Gamma^\nu(x,\xi)\phi(\xi)dV_\xi$$

is simply

$$[(n-2)\omega_n]^{-1}\{-\sum g_{ij}\partial_i\partial_j\int_{U_1}\phi_J(\xi)r(x,\xi)^{2-n}dV_\xi + \int_{U_1}\sum g_{ij}\partial_i\partial_j r(x,\xi)^{2-n}\phi_J(\xi)dV_\xi\}$$
$$+\sum_{k,l}\{\partial_k\partial_l\int_{U_1}D(x,\xi)dV_\xi - \int_{U_1}\partial_k\partial_l D(x,\xi)dV_\xi\}$$
$$+\sum_m\{\partial_m\int_{U_1}C(x,\xi)dV_\xi - \int_{U_1}\partial_m C(x,\xi)dV_\xi\}$$

where D and C are such that $\partial_k\partial_l D(x,\xi)$ and $\partial_m C(x,\xi)$ are absolutely summable over U and boundedly for $x\in U_1$—hence, by Lemma 0, the terms in $\{\ \}$'s vanish and we have that the J-th component of

$$(5.2) \qquad \nabla_x\left[\int_{U_1}\Gamma^\nu(x,\xi)\phi(\xi)dV_\xi\right] - \int_{U_1}\nabla_x\Gamma^\nu(x,\xi)\phi(\xi)dV_\xi$$

is

$$(5.3) \quad -\sum g_{ij}(x)\Big[\partial_i\partial_j\int_{U_1}\{[(n-2)\omega_n r(x,\xi)^{n-2}]^{-1}\phi(\xi)dV_\xi\}$$
$$-\int_{U_1}\partial_i\partial_j\{[(n-2)\omega_n r(x,\xi)^{n-2}]^{-1}\phi(\xi)\}dV_\xi\Big].$$

Hence it is sufficient to prove that the latter expression is equal to $\phi_J(x)$, and this just amounts to proving Theorem A_0. We remark further that (5.3) is just what the whole expression (5.2) would equal if ∇ were the Laplacian $\Delta = d\delta + \delta d$, $\mathfrak{a}$ were the metric induced by the given invariant Riemannian metric, and ϕ_J were the sole component of a C^∞ differentiable form of degree 0. Hence it is sufficient to prove Theorem A_0 when $\nabla = \Delta$. Therefore it is sufficient to prove Theorem A^*_0 if $\nabla = \Delta$. But since $\Delta = d\delta + \delta d$ is self-adjoint, Theorem A^*_0 becomes:

$$(5.4) \qquad \eta(\xi) = \int_U [\Gamma^\nu(x,\xi)(\Delta\eta)(x) - (\Delta_x\Gamma^\nu(x,\xi))\eta(x)]\,dV_x, \quad \xi \in U_1,$$

because $\mathfrak{a}(\xi) = 1$, Δ is real, i.e., $\bar{\Delta} = \Delta$, and $\Gamma^\nu(x,\xi)$ is 1×1. (Though our operators Γ^ν and Q^ν do not coincide with the operators Ω and Q' of de Rham, our equation (5.4) is practically a special case of Lemma 3(I) p. 146 of [8] as far as the terms of significant contribution are concerned, and the calculations involved in the proof of Lemma 3, pp. 146-148 are practically the same as our calculations in proving (5.4).)

8. *Proof of equation* (5.4). If η, Γ are functions (0-forms), $\delta\eta = 0$, $\delta\Gamma = 0$. Hence

$$\Gamma^\nu(x,\xi)\Delta\eta(x) - (\Delta_x\Gamma^\nu(x,\xi))\eta(x) = \Gamma^\nu(x,\xi)\delta d\eta(x) - (\delta d\Gamma^\nu(x,\xi))\eta(x).$$

On the other hand, for 0-forms Γ and η,

$$\begin{aligned} d*(\Gamma d\eta - \eta d\Gamma) &= d(\Gamma * d\eta - \eta * d\Gamma) \\ &= d\Gamma \wedge *d\eta + \Gamma d*d\eta - d\eta \wedge *d\Gamma - \eta d*d\Gamma \\ &= \Gamma d*d\eta - \eta d*d\Gamma = (\eta\delta d\Gamma - \Gamma\delta d\eta). \end{aligned}$$

Hence it is enough to prove that

$$\eta(\xi) = -\int_U d*(\Gamma d\eta - \eta d\Gamma), \quad \xi \in U_1,$$

where $\Gamma = \Gamma^\nu(x,\xi)$, $\eta = \eta(x)$, and d and $*$ are "with respect to the variable x."

We introduce normal coordinates ζ^i with center ξ such that in these coordinates our given metric $\{g_{ij}\}$ satisfies $g_{ij}(\xi) = \delta_{ij}$. Then the geodesic distance from ξ of a point with normal coordinates $\zeta^1, \cdots, \zeta^n$ is $\sum_{i=1}^{n} (\zeta^i)^2$. For small $\alpha > 0$ we let $\sigma_\alpha = \{\zeta \mid \sum(\zeta^i)^2 \leqq \alpha^2\}$ be the geodesic sphere of radius α with center ξ. Moreover, we let S_α be the surface of σ_α, $S_\alpha = \{\zeta \mid \sum(\zeta^i)^2 = \alpha^2\}$, and let $U_\alpha = U - \sigma_\alpha$. Then since $\mathfrak{L}(\eta) \subset U$ and $\mathfrak{L}(\eta)$ is compact, we may apply Stokes' Theorem to obtain

$$\int_U d*(\Gamma d\eta - \eta d\Gamma) = \lim_{\alpha\to 0} \int_{U_\alpha} d*(\Gamma d\eta - \eta d\Gamma) = \lim_{\alpha\to 0} \int_{S_\alpha} *(\Gamma d\eta - \eta d\Gamma).$$

Since the $(n-1)$-dimensional "area" of S_α is $0(\alpha^{n-1})$, and since $|\Gamma| \leqq C\alpha^{2-n}$ on S_α, $\lim_{\alpha\to 0}\int_{S_\alpha} *(\Gamma d\eta) = 0$. Therefore

$$\int_U d*(\Gamma d\eta - \eta d\Gamma) = \lim_{\alpha\to 0}\int_{S_\alpha} *(\eta d\Gamma).$$

Moreover, $\Gamma^\nu(x,\xi) = [(n-2)\omega_n]^{-1} r(x,\xi)^{2-n} + A(x,\xi)$, where $r(x,\xi)^{n-2}A(x,\xi)$ approaches zero as $r(x,\xi)$ approaches zero, so that

$$\lim_{\alpha\to 0}\int_{S_\alpha} *\eta d\Gamma = \lim_{\alpha\to 0}[(n-2)\omega_n]^{-1}\int_{S_\alpha} \eta * d(r(x,\xi)^{2-n}).$$

$\eta \in C^\infty$, and therefore on S_α $\eta = \eta(\xi) + O(\alpha)$. Hence

$$\lim_{\alpha\to 0}\int_{S_\alpha} *\eta d\Gamma = \eta(\xi)[(n-2)\omega_n]^{-1}\lim_{\alpha\to 0}\int_{S_\alpha} *d(r(x,\xi)^{2-n}).$$

We re-employ our previous notation, $P(x,\xi) = r(x,\xi)^2$, $\tau = (n-2)/2$ (the treatment is similar if $n = 2$). Then

$$\int_{S_\alpha} *d(r(x,\xi)^{2-n}) = \int_{S_\alpha} *dP^{-\tau} = -\int_{S_\alpha} *\tau P^{-\tau-1}dP = -\tau\alpha^{-n}\int_{S_\alpha} *dP,$$

since $P(x,\xi) = \alpha^2$ for $x \in S_\alpha$. However, $P \in C^\infty$, and hence we may apply Stokes' theorem again obtaining

$$\int_{S_\alpha} *d(r(x,\xi)^{2-n}) = -\tau\alpha^{-n}\int_{\sigma_\alpha} d*dP = \tau\alpha^{-n}\int_{\sigma_\alpha} *\delta dP$$
$$= \tau\alpha^{-n}\int_{\sigma_\alpha} *\Delta P = \tau\alpha^{-n}\int_{\sigma_\alpha} \Delta P*1,$$

where, in the given system of coordinates, $*1 = \mathfrak{g}^{\frac{1}{2}} d\zeta^1 \wedge \cdots \wedge d\zeta^n$ and $\mathfrak{g}^{\frac{1}{2}} = (\det(\mathfrak{g}_{ij}))^{\frac{1}{2}}$. Referring to formulas 3(ix), 3(x), and (4.9) with $W = 1$ and $l(P) = P$, we see that $\Delta P = 2n - 2\mathfrak{R}(x,\xi)$, where $|\mathfrak{R}(x,\xi)| \leqq K\alpha$ for $x \in \sigma_\alpha$, K being a constant independent of α. Therefore

$$\lim_{\alpha\to 0}\tau\alpha^{-n}\int_{\sigma_\alpha} \Delta P*1 = \lim_{\alpha\to 0} n(n-2)\alpha^{-n}\int_{\sigma_\alpha} *1.$$

Since $\sqrt{\mathfrak{g}(\xi)} = 1$, $\sqrt{\mathfrak{g}(x)} = 1 + O(\alpha)$ for $x \in \sigma_\alpha$, so that

$$\lim_{\alpha\to 0}\alpha^{-n}\int_{\sigma_\alpha} *1 = \lim_{\alpha\to 0}\alpha^{-n}\cdot\frac{\omega_n}{n}\alpha^n(1 + O(\alpha)) = \frac{\omega_n}{n}.$$

Therefore

$$\lim_{\alpha\to 0}\int_{S_\alpha} *(\eta d\Gamma) = \eta(\xi)[(n-2)\omega_n]^{-1}\cdot(n-2)n\cdot\frac{\omega_n}{n} = \eta(\xi),$$

which is our desired result.

We can now prove the regularity theorem:

THEOREM W. *Let ∇ be a strongly elliptic operator as above. Let $\phi \in \mathfrak{J}$ and suppose that for all $\psi \in \mathfrak{J}_\infty$ such that the carrier of ψ is compact and contained in U we have*

$$(w) \qquad \int_U {}^t\phi \mathfrak{A} \overline{\nabla^* \psi} dV = 0.$$

Then $\phi \in \mathfrak{J}_\infty[U_1]$ mod $\mathfrak{N}$ *and $\nabla\phi = 0$.*

Suppose, moreover, that all $\mathfrak{M}_g$ commute with ∇ and ∇^*. Then if $\phi \in \mathfrak{J}$ is G-invariant and if (w) holds for all G-invariant $\psi \in \mathfrak{J}_\infty$ such that the carrier of ψ is compact and contained in U, we still have $\phi \in \mathfrak{J}_\infty[U_1]$ and $\nabla\phi = 0$.

Proof. Let $\eta \in \mathfrak{J}_\infty[U_1]$ be such that the carrier of η is compact and contained in U_1. Let

$$\psi(x) = \int_U \Gamma^\nu(x,\xi)\eta(\xi) dV_\xi.$$

Then $\mathfrak{L}(\psi)$ is compact and contained in U, and by Lemma β, $\psi \in C^\infty$. Therefore, by Theorem A_N,

$$0 = \int_U {}^t\phi \mathfrak{A} \overline{\nabla^* \psi}\, dV = \int_U {}^t\phi(x) \mathfrak{A}(x) [\overline{\eta(x)} + \int_{U_1} \overline{Q^\nu(x,\xi)\eta(\xi)} dV_\xi] dV_x,$$

or, since $\mathfrak{L}(\eta)$ is contained in U_1,

$$\int_{U_1} {}^t\phi \mathfrak{A} \bar{\eta}\, dV_\xi = -\int_{U_1} (\int_U {}^t\phi(x) \mathfrak{A}(x) \overline{Q^\nu(x,\xi)} dV_x) \overline{\eta(\xi)} dV_\xi.$$

This holds for all $\eta \in \mathfrak{J}_\infty[U_1]$ such that $\mathfrak{L}(\eta)$ is compact and contained in U_1. This shows that on U_1

$${}^t\phi(\xi) = -\int_U {}^t\phi(x) \mathfrak{A}(x) \overline{Q^\nu(x,\xi)} dV_x, \quad \operatorname{mod} \mathfrak{N};$$

but it follows from Lemma δ that the right side is of class C^ν; since ν is arbitrary, we conclude that $\phi \in C^\infty \operatorname{mod} \mathfrak{N}$. Then by the definition of ∇^* it follows immediately that $\nabla\phi = 0$. This concludes the proof of the first part of the theorem.

Now suppose that ϕ is G-invariant and that for all G-invariant $\psi \in \mathfrak{J}_\infty$, $\mathfrak{L}(\psi) \subset U$, we have $\int_U {}^t\phi \mathfrak{A} \nabla^* \phi\, dV = 0$. Let $\theta \in \mathfrak{J}_\infty$, $\mathfrak{L}(\theta) \subset U$. Since our chosen Riemannian metric is G-invariant, since our chosen metric for the bundle is G-invariant, and since $\mathfrak{M}_g$ commutes with ∇ and ∇^* for each $g \in G$, we have, because ϕ is G-invariant,

$$\int_U {}^t(\phi)\,\mathcal{A}\,(\overline{\nabla^*\theta})\,dV = \int_U {}^t(\phi_g)\,\mathcal{A}\,(\overline{\nabla^*\theta})\,dV$$
$$= \int_U {}^t\phi\mathcal{A}\,\overline{(\nabla^*\theta)}_g dV = \int_U {}^t\phi\mathcal{A}\,\overline{\nabla^*(\theta_g)}\,dV,$$

where θ_g has the same meaning as at the beginning of Section 4. Therefore, if we put $\theta^* = (\mathrm{ord}\, G)^{-1} \sum_g \theta_g$, θ^* is G-invariant and

$$\int_U {}^t\phi\mathcal{A}\,\overline{\nabla^*\theta}\,dV = \int_U {}^t\phi\mathcal{A}\,\overline{\nabla^*\theta^*}\,dV = 0$$

by hypothesis. Therefore, by the first part of the theorem, $\phi \in C^\infty$, and $\nabla\phi = 0$. Q. E. D.

6. The decomposition theorem for compact V-manifolds. Let $\mathcal{V}$ be a compact C^∞ V-manifold of real dimension n, and let $\mathcal{V}$ be supplied with a C^∞ Riemannian metric g. Let dV denote the measure on $\mathcal{V}$ associated with g. We let B be a C^∞ V-bundle over $\mathcal{V}$ having a N-dimensional vector space as fibre, and let $\mathfrak{a}$ be a C^∞ metric for B. Finally, denote by M the module of C^∞ sections of B. We return to our previous assumption that for each $\{U, G, \varphi\} \in \mathcal{F}$, B_U is the product bundle. Then we employ the notation of the preceding section and in a given l. u. s. let $h_U(g)^{-1} = \mathcal{M}_g$, denote $\mathfrak{a}_U$ simply by $\mathcal{A}$, and dV_U by dV. If ∇ is an endomorphism of M, we say that ∇ is strongly elliptic if (4. 2) holds in each l. u. s. $\{U, G, \varphi\}$, it being understood, by definition, that ∇ commutes with the action of $\mathcal{M}_g$ on B_U for each $g \in G$. We let ∇ be a strongly elliptic endomorphism of M for the rest of this section. Then ∇ has a strongly elliptic adjoint ∇^* which satisfies $(\phi, \nabla^*\psi) = (\nabla\phi, \psi)$.

We let $\mathcal{J}_{\mathcal{V}}$ denote the space of square-summable measurable section of B supplied with inner product of Section 1. We let $\mathcal{H}$ and $\mathcal{H}^*$ denote respectively the null spaces of ∇ and ∇^* in $M \subset \mathcal{J}_{\mathcal{V}}$. We first observe that $\mathcal{H}$ and $\mathcal{H}^*$ are closed in $\mathcal{J}_{\mathcal{V}}$. This is an immediate consequence of Theorem W. We then let $\mathcal{L}$ and $\mathcal{L}^*$ denote respectively the orthogonal complements of $\mathcal{H}$ and $\mathcal{H}^*$ in $\mathcal{J}_{\mathcal{V}}$ and define $\mathcal{L}_\infty = \mathcal{L} \cap M$, $\mathcal{L}_\infty{}^* = \mathcal{L}^* \cap M$. Clearly $M = \mathcal{L}_\infty + \mathcal{H} = \mathcal{L}_\infty{}^* + \mathcal{H}^*$. We first prove

LEMMA M. (*Minimum Eigenvalue*). *There exists $c > 0$ such that for all $\phi \in \mathcal{L}_\infty$, $\phi \neq 0$, $\| \nabla\phi \| / \| \phi \| \geqq c$.*

Proof. If the lemma were false, we could find a sequence $\{\phi_m\}$, $\phi_m \in \mathcal{L}_\infty$, such that $\| \phi_m \| = 1$ and $\lim_{m\to\infty} \| \nabla\phi_m \| = 0$. At each $x \in \mathcal{V}$ we can find a l. u. s. $\{U(x), G, \varphi\}$ and $U_1(x) \subset U(x)$ such that $U(x)$ and $U_1(x)$ enjoy respectively the same properties as U and U_1 of Section 5. We let

$U_1(x^1), \cdots, U_1(x^k)$ cover $\mathfrak{V}$ and denote $U(x^i)$ and $U_1(x^i)$ respectively U^i and $U_1{}^i$, $i=1,\cdots,k$. Moreover, for U^i and $U_1{}^i$ we let $\Gamma_i{}^\nu$ and $Q_i{}^\nu$ be the integral operators defined previously for any integer $\nu > 0$. Since $\Gamma_i{}^\nu$ is a bounded operator,

$$(6.1)\quad \lim_{m\to\infty} \| \phi_m + Q_i{}^\nu \phi_m \|_{U_1{}^i} = \lim_{m_j\to\infty} \| \Gamma_i{}^\nu \nabla \phi_m \|_{U_1{}^i} \leqq \lim_{m\to\infty} K_i \| \nabla \phi_m \|_{U_1{}^i} = 0, \quad i=1,\cdots,k,$$

and hence $\lim\limits_{m\to\infty} \| \phi_m + Q_i{}^\nu \phi_m \| = 0$, while on the other hand, since $Q_i{}^\nu$ is completely continuous, we can, by a diagonalization process, find a subsequence $\{\phi_{m_j}\}$ of $\{\phi_m\}$ such that for each i, $\lim\limits_{m_j\to\infty} \| Q_i{}^\nu \phi_{m_j} - \mathrm{X}_i \|_{U_1{}^i} = 0$, where $\mathrm{X}_i \in \mathfrak{J}[U_1{}^i]$. Then by (6.1), $\phi_{m_j} \to -\mathrm{X}_i$, on $U_1{}^i$ and we may take X_i to be G^i-invariant since the subspace of G^i-invariant elements of $\mathfrak{J}[U_1{}^i]$ is closed. It is seen that on $U_1{}^i \cap U_1{}^k$, $\mathrm{X}_i \equiv \mathrm{X}_k \bmod \mathfrak{N}$, and hence ϕ defined by $\phi \equiv -\mathrm{X}_i \bmod \mathfrak{N}$ in $U_1{}^i$ is a well-defined measurable section of B, $\lim \| \phi_{m_j} - \phi \| = 0$, and $\| \phi \| = 1$ since $\| \phi_m \| = 1$. On the other hand, since $\lim \| \nabla \phi_m \| = 0$, and $|(\nabla^* \eta, \phi_{m_j})| = |(\nabla \phi_{m_j}, \eta)| \leqq \| \nabla \phi_{m_j} \| \cdot \| \eta \|$ for each $\eta \in M$, it follows that $(\nabla^* \eta, \phi) = 0$ for all $\eta \in M$, and hence, by Theorem W, $\phi \in C^\infty$ and $\nabla \phi = 0$, i.e., $\phi \in \mathfrak{H}$. But then $(\phi_m, \phi) = 0$ for each m and therefore $\| \phi \| = 0$, which contradicts $\| \phi \| = 1$. Therefore our lemma is established.

If ∇ is self-adjoint, i.e., $\nabla = \nabla^*$, then $\mathfrak{H} = \mathfrak{H}^*$ and $\mathfrak{L} = \mathfrak{L}^*$. We can now prove

THEOREM B. *Assume that ∇ is self-adjoint. Then for each $\beta \in \mathfrak{L}_\infty$ there is a unique $\phi \in \mathfrak{L}_\infty$ such that $\nabla \phi = \beta$.*

Proof. Since $\mathfrak{H}$ is the null-space of $\nabla = \nabla^*$, it is clear that $\nabla \mathfrak{L}_\infty \subset \mathfrak{L}_\infty$.

Moreover, $\nabla \mathfrak{L}_\infty$ is everywhere dense in $\mathfrak{L}$. To see this, let $[\nabla \mathfrak{L}_\infty]$ be the closure of $\nabla \mathfrak{L}_\infty$ in $\mathfrak{L}$ and let $\mathfrak{M}$ be its orthogonal complement, $\mathfrak{L} = [\nabla \mathfrak{A}_\infty] + \mathfrak{M}$. Let $m \in \mathfrak{M}$. Then $m \perp \nabla \mathfrak{L}_\infty$. Since $M = \mathfrak{L}_\infty + \mathfrak{H}$, $\nabla M = \nabla \mathfrak{L}_\infty$ and hence $m \perp \nabla M$. Therefore, by Theorem W, m is of class C^∞ and $\nabla m = 0$; thus $m \in \mathfrak{H}$. However, $m \in \mathfrak{M} \subset \mathfrak{L} \perp \mathfrak{H}$. Hence $m = 0$. Therefore $[\nabla \mathfrak{L}_\infty] = \mathfrak{L}$.

Now let $\beta \in \mathfrak{L}_\infty \subset \mathfrak{L}$. Then there exists a sequence $\{\phi_m\}$, $\phi_m \in \mathfrak{L}_\infty$, such that $\nabla \phi_m \to \beta$. Hence $\nabla \phi_m$ is a Cauchy sequence and thus, by Lemma M, $\{\phi_m\}$ is itself a Cauchy sequence with limit $\phi \in \mathfrak{L}$. In each $U_1{}^i$,

$$\Gamma_i{}^\nu \nabla \phi_m = \phi_m + Q_i{}^\nu \phi_m.$$

We know by Lemma ϵ that $\| \Gamma_i{}^\nu \nabla \phi_m - \Gamma_i{}^\nu \beta \|_{U_1{}^i} \leqq C_i \| \nabla \phi_m - \beta \|_{U^i}$, while since $Q_i{}^\nu$ is completely continuous it is automatically bounded and therefore

$\| Q_i{}^\nu \phi_m - Q_i{}^\nu \phi \| \leqq C_i{}^* \| \phi_m - \phi \|$. Therefore $\Gamma_i{}^\nu \beta \equiv \phi + Q_i{}^\nu \phi \mod \mathfrak{N}$. By Lemma α, $\Gamma_i{}^\nu \beta$ is of class C^∞ and by Lemma δ, $Q_i{}^\nu \phi$ is of class C^ν. Since this is true for each positive integer ν and all i, $\phi \in C^\infty \mod \mathfrak{N}$; thus $\phi \in \mathcal{L} \mod \mathfrak{N}$. Moreover $\nabla \phi \equiv \beta \mod \mathfrak{N}$; for if $\eta \in M$,

$$(\nabla \phi_m, \eta) = (\phi_m, \nabla \eta), \qquad (\nabla \phi_m, \eta) \to (\beta, \eta),$$

and $(\phi_m, \nabla \eta) \to (\phi, \nabla \eta) = (\nabla \phi, \eta)$. Hence $(\beta - \nabla \phi, \eta) = 0$ for all $\eta \in M$ and therefore $\beta - \nabla \phi = 0$.

If $\nabla \phi' = \beta$, $\| \phi - \phi' \| \leqq C \| \nabla \phi - \nabla \phi' \| = 0$ and hence $\phi \equiv \phi' \mod \mathfrak{N}$. Therefore ϕ is unique and our proof is complete.

We have as a corollary

THEOREM D (*Decomposition Theorem*). *If ∇ is self-adjoint,*

$$M = \nabla M \oplus \mathcal{H}.$$

Proof. By what we have just proved, $\nabla M = \mathcal{L}_\infty$, and by construction $M = \mathcal{L}_\infty \oplus \mathcal{H}$. Hence $M = \nabla M \oplus \mathcal{H}$. q. e. d.

THEOREM F. *$\mathcal{H}$ is finite dimensional.*

Proof. It is sufficient to prove that we can extract a Cauchy sequence from every bounded sequence in $\mathcal{H}$. Let U^i, $U_1{}^i$, $\Gamma_i{}^\nu$, and $Q_i{}^\nu$ have the same meanings as in the proof of Lemma M. Let $\{\phi_m\}$ be a bounded sequence in $\mathcal{H}$. We have

$$0 = \Gamma_i{}^\nu \nabla \phi_m = \phi_m + Q_i{}^\nu \phi_m, \text{ on } U_1{}^i,$$

for each i and every ϕ_m. Since $\{\phi_m\}$ is a bounded sequence and $Q_i{}^\nu$ is completely continuous, we can find by diagonalization a subsequence $\{\phi_{m_j}\}$ such that for each i $\{Q_i{}^\nu \phi_{m_j}\}$ is a Cauchy sequence on $U_1{}^i$ under the norm $\| \; \|_{U_1{}^i}$. Since $\| \phi_{m_j} - \phi_{m_k} \| \leqq \sum_i \| \phi_{m_j} - \phi_{m_k} \|_{U_1{}^i} = \sum_i \| Q_i{}^\nu \phi_{m_j} - Q_i{}^\nu \phi_{m_k} \|_{U_1{}^i}$, $\{\phi_{m_j}\}$ is itself a Cauchy sequence, and our proof is complete.

7. Hodge's and Kodaira's theorems for V-manifolds. Let $\mathcal{V}$ be a compact V-manifold supplied with a C^∞ Riemannian metric g. Let A^k be the V-bundle of differential forms of degree k over $\mathcal{V}$ (see Section 1). Then A^k is a C^∞ V-bundle over $\mathcal{V}$ with a vector space of dimension $\binom{n}{k}$ as fibre, and the Laplacian $\Delta = d\delta + \delta d$ is a self-adjoint, strongly elliptic endomorphism of the module M^k of C^∞ sections of A^k. Denote respectively by $\mathcal{H}^k$ and Z^k the null-spaces of Δ and d in M^k. Then by Theorem D, $M^k = \Delta M^k + \mathcal{H}^k$. It is clear that the spaces $d\delta M$ and $\delta d M$ are orthogonal since $d^2 = 0$ and $\delta^2 = 0$. Therefore $M^k = d\delta M^k \oplus \delta d M^k \oplus \mathcal{H}^k$. Since

$$(\phi, \Delta\phi) = (d\phi, d\phi) + (\delta\phi, \delta\phi),$$

it is clear that if $\phi \in \mathfrak{H}^k$, then $d\phi = 0$. Suppose now that $\phi \in M^k$, $\phi = d\delta a + \delta db + h$, $d\phi = 0$; then $d\delta db = 0$, which shows that $\delta db = 0$. Moreover, it is clear that $d\delta M^k = dM^{k-1}$. Hence $\mathfrak{H}^k = Z^k/dM^{k-1}$. On the other hand, Satake has shown that $H^k(\mathfrak{V}, R) = Z^k/dM^{k-1}$ where $H^k(\mathfrak{V}, R)$ is the k-dimensional Čech cohomology group of $\mathfrak{V}$ with real coefficients.

THEOREM H (*Hodge's Theorem*). $H^k(\mathfrak{V}, R) \cong \mathfrak{H}^k$.

Let B be a complex line bundle over $\mathfrak{V}$, and let B be supplied with a C^∞ metric a. Denote by $M(B)^{r,s}$ the module of C^∞ sections of $B \otimes A^{r,s}$, and denote by $\mathfrak{H}(B)^{r,s}$ and $Z(B)^{r,s}$ respectively the null-spaces of $\square$ and $\bar{\partial}$ in $M(B)^{r,s}$. Then reasoning parallel to that above (see Kodaira, [5]) yields

THEOREM K (*Kodaira's Theorem*). $\mathfrak{H}(B)^{r,s} = Z(B)^{r,s}/\bar{\partial}M(B)^{r,s-1}$. This will lead to a natural generalization of Dolbealut-Kodaira's Theorem (see [5]) for V-manifolds, which we shall prove in a later paper, and which we shall in turn use to prove an imbedding theorem which slightly generalizes that of [1].

PRINCETON UNIVERSITY.

REFERENCES.

[1] W. L. Baily, "On the quotient of an analytic manifold . . . ," *Proceedings of the National Academy of Sciences*, vol. 40 (1954), pp. 804-808.

[2] S. Banach, *Theorie des Operations Lineaires*, Warsaw, 1932.

[3] J. Hadamard, *Le Problem de Cauchy et les équations aux dérivées partielles linéaires hyperboliques*, Paris, 1932.

[4] K. Kodaira, "Harmonic fields in Riemannian manifolds," *Annals of Mathematics*, vol. 50 (1949), pp. 587-665.

[5] ———, "On cohomology groups of compact analytic varieties with coefficients in some analytic faisceaux," *Proceedings of the National Academy of Sciences*, vol. 39 (1953), pp. 865-868.

[6] F. J. Murray and K. S. Miller, *Existence Theorems for Ordinary Differential Equations*, New York, 1954.

[7] S. Nakano, "On complex analytic vector bundles," *Journal of the Mathematical Society of Japan*, vol. 7 (1955), pp. 1-12.

[8] G. deRham, *Varietes Differentiables*, Paris, 1955.

[9] I. Satake, "On a generalization of the notion of Manifolds," *Proceedings of the National Academy of Sciences*, vol. 42 (1956).

RIEMANN METRICS ASSOCIATED WITH CONVEX BODIES AND NORMED SPACES.*

By Detlef Laugwitz and Edgar R. Lorch.

1. Introduction. It is the aim of this paper to develop a connection between convex bodies in normed spaces (or Minkowski spaces) and Riemann differential geometry.

We shall use the following notations of tensor calculus with the extension to infinite-dimensional spaces as introduced in [3]. Vectors of the Banach space $\mathfrak{B}^1$ will be denoted by x, y, z, and these letters may bear a superscript: x^i, y^k, etc. Let $\mathfrak{B}_1$ denote the adjoint space of $\mathfrak{B}^1$ (space of covariant vectors). We shall denote the elements of $\mathfrak{B}_1$ by $\xi, \eta, \cdots$, or also by these same letters with a subscript. We shall be concerned only with reflexive spaces $\mathfrak{B}^1$, for which the definition of a tensor of higher order is simply the following. Let us consider bounded multilinear functionals defined upon $\mathfrak{B}^1 \times \cdots \times \mathfrak{B}^1$ (n times) and $\mathfrak{B}_1 \times \cdots \times \mathfrak{B}_1$ (m times). Such a functional will be denoted by $t_{k_1 \cdots k_n}{}^{i_1 \cdots i_m}$. The Banach spaces $\mathfrak{B}_n{}^m$ constituted by these (n, m)-multilinear functionals permit a generalization of the classical summation convention: If in an expression the letter i, say, occurs twice, once as subscript of one functional, and once as superscript of another one, the associated linear operation is to be effected. For instance, $t_k{}^i x^k \eta_i$ stands for the numerical value of the bilinear functional t for $x \in \mathfrak{B}^1$, $\eta \in \mathfrak{B}_1$. For a detailed statement the reader is referred to the paper [3].

For finite-dimensional spaces this notation may be considered in the manner of classical tensor calculus after the choice of any special base. But it should be noted that our interpretation permits a direct operation with vectors and tensors and not only with their components with respect to a base. This point of view seems to be useful for the study of vector spaces. There is no longer any need for proofs of invariance of formulas under a change of base.

These conventions apply to Fréchet differentials. For instance, the derivative of a scalar function $f(x)$ is a bounded linear functional and may be written $\partial f(x) / \partial x^i$.

* Received April 26, 1956.

2. The Riemann metric in $\mathfrak{B}^1$ associated with a convex body $\mathcal{B}$. We consider a bounded convex body $\mathcal{B}$ in Hilbert space $\mathfrak{B}^1$ containing the origin 0 as an inner point. It is well known that with $\mathcal{B}$ a norm or Minkowski metric $F(x)$ is associated which is continuous, positively homogeneous of order 1, and convex, and such that $\mathcal{B}$ coincides with the point set $F(x) \leqq 1$. We shall suppose that $\mathcal{B}$ is sufficiently smooth, precisely, that $g(x) = \frac{1}{2}F^2(x)$ has Fréchet differentials (for $x \neq 0$) up to the fifth order.

Denoting the Fréchet derivatives of g by subscripts, we have $g_i \in \mathfrak{B}_1$, $g_{ik} \in \mathfrak{B}_2$, $g_{ijk} \in \mathfrak{B}_3$, etc. We shall now impose a last condition on $\mathcal{B}$:

$$g_{ik}(x) y^i y^k \geqq m \cdot \| y \|^2 \text{ for } y \neq 0 \text{ with } m = m(x) > 0. \tag{1}$$

This condition implies convexity of $\mathcal{B}$ in the usual sense [6, § 5]. The functional $g_{ik}(x)$ is homogeneous of order zero, and

$$g_{ik}(x) x^i x^k = 2g(x) = F^2(x). \tag{2}$$

From Euler's theorem on homogeneous functions we have

$$g_{ijk}(x) x^k = 0. \tag{3}$$

It has been stated previously [5, p. 107], that g_{ik} is not only the fundamental tensor of the Minkowski metric $F(x)$, but moreover defines a Riemann metric throughout $\mathfrak{B}^1$ (with the exception of the origin 0):

$$ds^2 = g_{ik}(x)\, dx^i dx^k. \tag{4}$$

In the present paper the authors begin the study of this Riemann metric defined by $\mathcal{B}$.

3. The curvature tensor. We shall denote by $g^{ij}(x)$ the inverse tensor of $g_{ik}(x)$,

$$g^{ij}(x) g_{jk}(x) = \delta_k{}^i, \tag{5}$$

where $\delta_k{}^i$ denotes the identity operator $\delta_k{}^i x^k = x^i$; g^{ij} exists because of (1) and [3, Satz 3.1]. The affine connection of the Riemann metric $g_{ik}(x)$ is represented by

$$\Gamma_{kj}{}^i = \tfrac{1}{2} g^{ir}(g_{rkj} - g_{kjr} + g_{jrk}) = \tfrac{1}{2} g^{ir} g_{rkj}. \tag{6}$$

From this we may compute the curvature tensor (definition for the infinite-dimensional case see [3, p. 137]):

$$R^l{}_{ijk} = \tfrac{1}{4} g^{lr} g^{ms}(g_{mlj} g_{rsk} - g_{mik} g_{rsj}). \tag{7}$$

This formula follows by the use of $g_{ijm}g^{jk}+g_{ij}\partial g^{jk}/\partial x^m=0$, which is obtained by differentiation from (5).

3.1. *For two-dimensional spaces $\mathfrak{B}^1$ the curvature tensor* (7) *of the Riemann metric $g_{ik}(x)$ associated with any body $\mathfrak{C}$ vanishes identically.*

Proof. Let x denote an arbitrary point, $x\neq 0$. We introduce a special base e_1, e_2 by $e_1=x$ and $g_{ik}(e_1)e_1{}^ie_2{}^k=0$, $g_{ik}(e_1)e_2{}^ie_2{}^k=1$. The components of $\bar{g}_{ijk}(e_1)$ with respect to this base are $\bar{g}_{111}(e_1)=g_{ijk}(e_1)e_1{}^ie_1{}^je_1{}^k$, etc. By Euler's theorem (3) we obtain

$$0=g_{ijk}(e_1)e_1{}^k=\bar{g}_{ij1}(e_1), \tag{8}$$

or, in other words, $\bar{g}_{ijk}(e_1)=0$ if at least one subscript equals 1. Now the only interesting component of the two-dimensional curvature tensor $R_{ijkl}=g_{ih}R^h{}_{jkl}$ is R_{1212}; but if $j=1$ the expression on the right hand side of (7) vanishes because of (8). From the tensor property of R we have $R^i{}_{jkl}=0$ in every coordinate system.

Nothing shall be said here about the value of the curvature tensor for higher dimensions except the following remark:

3.2. *If the metric g_{ik} is of constant curvature K, then $K=0$.*

Proof. $R^i{}_{jkl}(x)x^j=0$ from (3) and (7). By this the curvature scalar vanishes for any two-dimensional surface element passing through the direction x. This proves 3.2.

4. Applications to E. Cartan's theory of Finsler spaces. The object of E. Cartan's theory of Finsler spaces [1] is a fibre space with a n-dimensional manifold as its basis space and the manifold of directions in n-dimensional vector space as its fibre. (Our results shall hold good for the infinite-dimensional Finsler spaces introduced in [4].) We consider here only the local point of view, that is, the tangent space of one fixed point. For this special case, Cartan's metric is $g_{ik}(x)=\partial^2 g(x)/\partial x^i\partial x^k$, and the components of the euclidean connection are

$$C_j{}^i{}_k(x)=\tfrac{1}{2}g^{ir}(x)g_{rjk}(x).$$

(To simplify notations, we write x instead of Cartan's x', which is possible since we consider a fixed point of Finsler space.) Obviously these geometric objects have the same expressions as the corresponding ones, namely g_{ik} and $\Gamma_{jk}{}^i$, in Section 2 of our paper, if we take for $\mathfrak{C}$ the indicatrix of the local

Minkowski metric. A more interesting connection is the relation between the curvature tensor R_{ijkl} of formula (7) and Cartan's tensor S_{ijkl}

$$R = 2g \cdot S, \tag{9}$$

which follows directly from the analytic expressions (7) and [1, (XVI), p. 34]. From (9), (1), and our Theorem 3.2 follows at once:

4.1. *For two-dimensional Finsler spaces Cartan's tensor S_{ijkl} vanishes identically.*

The Landsberg angle $d\phi$ ([2]; métrique angulaire en un point, [1, p. 13]) of two directions x, $x + dx$ $(F(x) = F(x + dx) = 1)$ is defined by $d\phi^2 = g_{ik}(x) dx^i dx^k$. Its geometrical meaning in the Riemann geometry introduced here is: *$d\phi$ equals the Riemann length of the arc joining the end points of the two unit vectors.* (See also O. Varga [8])

5. Connections with results of Varga. Recently O. Varga [8] has studied other Riemann geometries associated with $\mathfrak{G}$. In our terminology we may state his point of view simply as follows. Varga introduces the tensor $\gamma_{\alpha\beta} = x_\alpha{}^i x_\beta{}^k g_{ik}$, $x_\alpha{}^i = \partial x^i / \partial u^\alpha$, where $x^i(u^\alpha)$ is a parametric representation of the hypersurface (sphere) $F(x) = 1/\sqrt{K} = \text{const.}$, $K > 0$. Obviously $\gamma_{\alpha\beta}$ may be interpreted as the Riemann metric induced in the sphere by the Riemann metric (3) of the embedding space. Hence Varga's results concerning the curvature of this sphere may easily be regained by standard argumentations on subspaces of Riemann spaces. (See [7], and for the infinite-dimensional case to which we may as well extend Varga's theorems, see [3, p. 143].) We apply the general definition to the calculation of the second fundamental form $l_{\alpha\beta}$ of the hypersurfaces:

$$\begin{aligned} l_{\alpha\beta} = x_\alpha{}^i x_\beta{}^k \nabla_k n_i &= x_\alpha{}^i x_\beta{}^k (\partial_k n_i - \Gamma_{ik}{}^r n_r) = - K x_\alpha{}^i x_\beta{}^k (g_{ir} \partial_k x^r) \\ &= - \sqrt{K} x_\alpha{}^i x_\beta{}^k g_{ik} = - \sqrt{K} g_{\alpha\beta}, \end{aligned}$$

where Ricci's lemma, equation (3), and the normal vector $n^i = x^i \cdot K$. $n_i = g_{ik} x^k$ have been used. From Gauss' equation [7, p. 242] we obtain for the curvature of the sphere

$$R_{\alpha\beta\gamma\delta} = x_\alpha{}^i x_\beta{}^j x_\gamma{}^k x_\delta{}^l R_{ijkl} - K(g_{\alpha\gamma} g_{\beta\delta} - g_{\alpha\delta} g_{\beta\gamma})$$

which implies Varga's theorems.

6. A theorem on geodesics. By (6) the differential equation for the geodesics of the Riemann metric (4) is

$$\ddot{x}^i + \tfrac{1}{2} g^{ir}(x) g_{rlm}(x) \dot{x}^l \dot{x}^m = 0. \tag{10}$$

Obviously the straight lines through the origin $x^i(t) = t \cdot y^i$ are geodesics. Now the following question arises: What are the conditions for $\mathfrak{E}$ (or g) which imply that all straight lines (not only those passing through 0) are geodesics? The answer is:

6.1. *The system of geodesics of* (4) *coincides with the system of all straight lines if and only if* $\mathfrak{E}$ *is an ellipsoid having* 0 *as its center, or, in other words,* $g_{ik} =$ const.

Proof. Obviously this condition is sufficient. We shall first prove, that it is necessary for $n = 2$. By a well known theorem [7, p. 287] a necessary and sufficient condition for the geodesics to be straight lines (in a linear coordinate system) is $\Gamma_{kj}{}^i = \delta_k{}^i \phi_j + \delta_j{}^i \phi_k$, or, for any base,

$$\Gamma_{22}{}^1 = \Gamma_{11}{}^2 = 2\Gamma_{12}{}^1 - \Gamma_{22}{}^2 = 2\Gamma_{12}{}^2 - \Gamma_{11}{}^1 = 0.$$

In our case (equation (6)) these four equations may be written

$$\begin{aligned} g^{11} g_{122} + g^{12} g_{222} &= 0 \\ g^{21} g_{111} + g^{22} g_{112} &= 0 \\ 2g^{11} g_{112} + g^{12} g_{122} - g^{22} g_{222} &= 0 \\ -g^{11} g_{111} + g^{12} g_{112} + 2g^{22} g_{122} &= 0. \end{aligned}$$

The determinant of this system of four linear equations for the four unknowns g_{111}, g_{112}, g_{122}, g_{222} is $[(g^{11}g^{22}) - (g^{12})^2]^2 > 0$. Hence the system admits only the solution $g_{ijk} = 0$. That proves 6.1 for the two-dimensional case. For spaces $\mathfrak{B}^1$ of any dimension greater than 2 with the straight lines as geodesics of (4) the argumentation runs as follows. The straight lines are then a fortiori geodesics with regard to the Riemann metric induced in a two dimensional plane passing through the origin which is again of the form considered here. By the above result every two-dimensional subspace is euclidean (or, the indicatrix in it is an ellipse), and from this the euclidean character of the space follows by a well known theorem. This completes the proof of 6.1.

MATHEMATISCHES FORSCHUNGSINSTITUT, OBERWOLFACH, SCHWARZWALD
AND
BARNARD COLLEGE, COLUMBIA UNIVERSITY, NEW YORK, N. Y.

REFERENCES.

[1] E. Cartan, *Les espaces de Finsler*, Paris, 1934.

[2] G. Landsberg, "Krümmungstheorie und Variationsrechnung," *Jahresbericht der Deutschen Mathematiker-Vereinigung*, vol. 16 (1907), pp. 547-557.

[3] D. Laugwitz, "Differentialgeometrie ohne Dimensionsaxiom I, II," *Mathematische Zeitschrift*, vol. 61 (1954), pp. 100-118 and 134-149.

[4] ———, "Grundlagen für die Geometrie der unendlichdimensionalen Finslerräume," *Annali di Matematica Pura ed Applicata*, ser. 4, vol. 41 (1956), pp. ———.

[5] E. R. Lorch, "A curvature study of convex bodies in Banach spaces," *ibid.*, ser. 4, vol. 34 (1953), pp. 105-112.

[6] ———, "Differentiable inequalities and the theory of convex bodies," *Transactions of the American Mathematical Society*, vol. 71 (1951), pp. 243-266.

[7] J. A. Schouten, *Ricci calculus*, 2nd ed., Berlin, 1954.

[8] O. Varga, "Die Krümmung der Eichfläche des Minkowskischen Raumes und die geometrische Deutung des einen Krümmungstensors des Finslerschen Raumes," *Abhandlungen aus dem Mathematischen Seminar der Universität* Hamburg, vol. 20 (1955), pp. 41-51.

ADDENDA TO THE PAPER ON BÔCHER'S THEOREM

(vol. 76 (1954), pp. 183-190).*

By Aurel Wintner.

If $(a_{ik})=a=a(t)$, where $0\leqq t<\infty$, is an n by n matrix of continuous functions $a_{ik}(t)$, let $a\in S$ mean that $a(t)$ has the following property: Corresponding to every constant vector c, there exists a solution $x=x(t)$ of $dx/dt=a(t)x$ satisfying $x(t)\rightarrow c$ as $t\rightarrow\infty$ (in view of the principle of superposition, there is just one $x(t)=x_c(t)$ belonging to the "initial condition" $x(\infty)=c$, if $a\in S$). If $a\in R$ means the conditional integrability of $a(t)$, that is, the existence of a finite limit, as $T\rightarrow\infty$, of the integral of $a(t)$ over $0\leqq t\leqq T$, and if $n>1$, then $a\in R$ is neither necessary nor sufficient for $a\in S$ (cf. the paper quoted in the title, which will be referred to as *loc. cit.*). On the other hand. if $a\in L$ means (as usual, with $L=L^1$) that $\gamma\in R$ holds for the norm, $\gamma=|a|$, of a also, then, according to a criterion which goes back to Bôcher, $a\in L$ is sufficient for $a\in S$ (for references to the implication $L\Rightarrow S$, and for generalizations of $L\Rightarrow S$ derived from $L\Rightarrow S$ itself, cf. *loc. cit.*).

There is, however, something unsatisfactory in the standard criterion, $L\Rightarrow S$. In fact, this criterion fails to contain the fact that, as readily seen from a quadrature, $a\in R$ is sufficient (and necessary) for $a\in S$ in the scalar case, $n=1$. It is therefore of interest that $L\Rightarrow S$ can be improved to the following criterion: *If $a_0(t)$ denotes the diagonal matrix the diagonal elements of which are those of $a(t)$, then $a_0\in R$ and $a-a_0\in L$ together are sufficient for $a\in S$.* It will be clear from the following proof that (and in which manner) this extended criterion can be combined with the criteria given *loc. cit.*

First, it follows from $L\Rightarrow S$ by a known application of the method of the variation of constants (cf. vol. 68 (1946), p. 200, of this Journal), that if $\beta=\beta(t)$ is any continuous matrix the real part of the trace of which has an indefinite integral posessing a lower bound ($>-\infty$), then $\beta\in S$ and $a-\beta\in L$ suffice for $a\in S$. Since the trace condition is certainly satisfied if $\beta\in R$, it follows, by choosing $\beta=a_0$, that the italicized assertion is proved

* Received September 26, 1956.

if it is ascertained that $a_0 \in S$. But $a_0 \in S$ follows from the assumption $a_0 \in R$. In fact, since $\beta \in R$ is sufficient (and incidentally, as mentioned above, necessary as well) for $\beta = S$ in the scalar case, it is clear that the same is true in the case of any diagonal matrix β.

* *
*

If a scalar function $f(t)$ is continuous for $0 \leqq t < \infty$, what kind of its "smallness" (for large t) will ensure that the scalar differential equation $x'' + \{1 + f(t)\}x = 0$ represents an "adiabatic" perturbation of harmonic oscillator $y'' + y = 0$? By this is meant the existence of two solutions, say $x = x_j(t)$, where $j = \pm 1$, satisfying $x_j(t) \sim e^{jti}$ and $x_j'(t) \sim jie^{jti}$ as $t \to \infty$. Let this property of the coefficient function $1 + f(t)$ be symbolized by $f \in P$. Corresponding to the preceding notations, $f \in R$ and $f \in L$ will refer to the integrability and the absolute integrability of f (on $0 \leqq t < \infty$) respectively.

It is well-known that, as a consequence of the standard criterion $L \Rightarrow S$ for an a, the criterion $L \Rightarrow P$ holds for an f (for a direct proof, cf. pp. 71-73 of the London (1931) edition of Weyl's Quantum Mechanics). But the assumption $f \in L$ is quite drastic. It can be greatly improved by having recourse to the refinements of $L \Rightarrow S$ given *loc. cit.* In fact, the following improvement of $L \Rightarrow P$ can be obtained: *For $f \in P$ it is sufficient that each of the three functions f_k, where $k = -1, 0, 1$ and $f_k(t) = f(t)e^{2kti}$, satisfy the following two conditions: $f_k \in R$ and $F_k(t) \in L$, where*

$$F_k(t) = f(t) \int_t^\infty f(s) e^{-2ksi} \, ds \qquad \left(\int_t^\infty = \lim_{T \to \infty} \int_t^T \right).$$

The involvement of $f_{\pm 1}(t) = f(t)e^{\pm 2ti}$ (besides $f_0(t) \sim f(t)$ itself), being a *resonance* restriction, is curious, since the problem concerns itself with a *homogeneous* linear differential equation (in this connection cf., however, O. Perron, Math. Ztschr., vol. 32 (1930), p. 705 and Satz 1).

The proof proceeds as follows: It was shown *loc. cit.* that $a \in S$ whenever $a \in R$ and $\gamma \in L$, where $\gamma(t)$ is either of the matrix products $a(t)a^0(t)$, $a^0(t)a(t)$ in which $a^0(t)$ denotes the integral (R) of $a(t)$ over the half-line (t, ∞). On the other hand, it follows by a known application of the method of the variation of constants that $f \in P$ is equivalent to $a \in S$, where $a = a_f$ denotes the following binary matrix:

$$a_f(t) = f(t) \begin{pmatrix} -uv & -v^2 \\ u^2 & uv \end{pmatrix}, \text{ where } u = \cos t,\ v = \sin t$$

(cf., e.g., vol. 69 (1947), pp. 262-263 of this Journal, where the corresponding relation is calculated for the generalization $y'' + g(t)y = 0$ of the present $y'' + y = 0$). But it is readily seen that, for this $\alpha(t) = \alpha_f(t)$, the two matrix conditions $\alpha \in R$, $\alpha\alpha^0 \in L$ can be reduced to three pairs $f_k \in R$, $F_k \in L$ of scalar conditions.

An instructive illustration results by choosing $f(t) = (\sin at)/t$, where a is a real constant $\neq 0$. Then the standard sufficient condition, $f \in L$, for $f \in P$ is violated for every a. But it is clear that all $3 + 3$ conditions $f_k \in R$, $F_k \in L$ (with $F_k(t) = O(t^{-2})$ as $t \to \infty$) are satisfied unless $a = \pm 2$. Consequently, *$f \in P$ if $f(t) = (\sin at)/t$, provided that $a \neq \pm 2$.* This is the more interesting as the proviso, taking care of the "resonance effect" referred to above, is known to be indispensable. Cf. vol. 69 (1947), p. 269 of this Journal, where $a = 2$, and, for a result which claims much less (in fact, just the boundedness of the solutions) than the last italicized result if $a \neq \pm 2$, G. Ascoli, Rend. Acc. Naz. Lincei, ser. 8, vol. 9 (1950), pp. 210-213. Ascoli's corresponding general statements (p. 131) are based on the assumption that boundedness is the same thing as asymptotic equivalence, an assumption which, however, is not in general satisfied.

It may finally be mentioned that if the $3 + 3$ conditions of the preceding general theorem are satisfied by an $f(t)$, then it is seen, by a repetition of the preceding proof for the asymptotic equivalence with $y'' + y = 0$, that *the general criterion remains true if the preceding $x'' + \{1 + f(t)\}x = 0$ is replaced by $x'' + f(t)x' + x = 0$.*

THE JOHNS HOPKINS UNIVERSITY.

CORRESPONDENCE.

A correspondent, who wishes to remain anonymous, writes as follows:

SIR,

In a paper which has been published by your Journal (October, 1949), W. L. Chow proved the following theorem:

Every closed analytic subvariety V of the projective space $\boldsymbol{P}_r(\boldsymbol{C})$ is algebraic.

I should like to point out that there is a very simple proof for this *if V is assumed to be free from singularities.*

It is no restriction to assume that V is connected; let n be its complex dimension. In the polynomial ring $\boldsymbol{C}[X_0, \cdots, X_r]$, the homogeneous polynomials which vanish on V generate a homogeneous ideal $\mathfrak{p}$. Since V is connected, $\mathfrak{p}$ is a prime ideal and defines therefore an irreducible algebraic variety W, which is the smallest one containing V; call m its dimension; we have $m \geqq n$. Every rational function f on W can be written as $f = P/Q$, where P, Q are homogeneous polynomials of the same degree and Q is not in $\mathfrak{p}$; therefore f induces a meromorphic function on V. This shows that the field L of rational functions on W is isomorphic to a subfield of the field K of meromorphic functions on V. It is well-known (cf., e.g., C. L. Siegel's proof, Göttingen Nachrichten 1955) that K has at most the degree of transcendency n; so we get $m \leqq n$, and hence $m = n$; therefore W has the same dimension as V. It is also well-known that W is analytically irreducible (this follows from the fact that the set of all simple points on W is connected, which is easily proved by induction on the dimension, using suitable hyperplane sections). Therefore $W = V$.

Yours, etc.

X. X. X.

CORRECTIONS TO "ENGEL RINGS AND A RESULT OF HERSTEIN AND KAPLANSKY." *

By M. P. DRAZIN.

I. N. Herstein has pointed out to the author that the proof of Theorem 2. 1 in the paper named above (this *Journal*, vol. 77 (1955), pp. 895-913) is invalidated by the use of an incorrect expression for the typical element d of the ideal generated by x^m (the theorem may nevertheless be true, but must now be regarded as unproven). This in turn invalidates the proofs given for Lemma 4. 1 (i) and Theorems 3. 1, 4. 1 and 6. 2 (i. e. all those results depending on the explicit use of the Koethe radical).

These deletions involve consequential modifications in Theorems 5. 1 and 6. 1. In Theorem 5. 1, we must disallow the maximal alternative, and the proof in the minimal case must be based on Theorem 4. 5 rather than on Theorem 4. 1. In Theorem 6. 1, the final clause "in particular, by Theorem 4. 1, R is commutator-nil" should be deleted. Also, in the proof of Theorem 5. 4, the second sentence must be based on A. S. Amitsur's paper [1] ("On rings with identities," *Journal of the London Mathematical Society,* vol. 30 (1955), pp. 464-470) rather than on Theorem 4. 1.

Further, as already stated (this *Journal*, vol. 78 (1956), p. 224), Theorem 6. 5 can no longer be upheld, and, in place of Theorem 6. 4, we can substantiate only the following weaker assertion:

If a division ring D has the weak K-property, i. e. if there correspond to each pair $x, y \in D$ integers $k = k(x,y)$, $m = m(x,y)$ and $n = n(x,y)$ such that $e_k(x^m, y^n) = 0$, then, if we modify n appropriately, we may always replace k by 1 while keeping m as given; if D has characteristic zero then in fact no modification of n is needed, i. e. $[x^{m(x,y)}, y^{n(x,y)}] = 0$ for all x, y.

It goes without saying that many of the linking passages and informal remarks in the paper under correction must now be looked on as (at best) misleading, but we shall not discuss these here. However, the remaining formally enunciated results, and in particular the key Theorems 4. 3, 4. 4 and 4. 5, are indeed true as originally stated.

TRINITY COLLEGE, CAMBRIDGE, ENGLAND.

* Received August 27, 1956.

ERRATA.

S. Abhyankar, "On the ramification of algebraic functions," this JOURNAL, vol. 77 (1955), pp. 575-592.

(1) Page 580 line 6 in the proof of Theorem 1: Instead of "P^* is at finite distance" read "$g_i \varepsilon M^*$ for $i = 1, 2, \cdots, m$."

(2) Page 581 line 6: Instead of "Y_{m-r}" read "Y_{m+r}."

(3) Page 581 lines 7 and 8: Instead of "$i = m+1, m+2, \cdots, m+r$," read "$i = 1, 2, \cdots, m$."

(4) Page 581 line 26; Instead of "G as a point in $S_{m+1}/k(x_1, x_2, \cdots, x_{r-1})$," read "$U^*$ as a point in $S_{m+1}/k(x_2, x_3, \cdots, x_r)$."

(5) Page 581 lines 27 and 28: Twice instead of "at $Y_1 = y_1$, $Y_2 = y_2$, $\cdots, Y_{m+r} = y_{m+r}$" read "at U^*."